

# Two-Step versus Simultaneous Estimation of Survey-Non-Sampling Error and True Value Components of Small Area Sample Estimators

P.A.V.B. Swamy<sup>a</sup>, T.S. Zimmerman<sup>b</sup> and J.S. Mehta<sup>c</sup>

<sup>a,b</sup> Bureau of Labor Statistics, Room 4985, 2 Massachusetts Avenue, N.E., Washington, DC 20212, E-mail: [paravastu\\_s@bls.gov](mailto:paravastu_s@bls.gov) and [Zimmerman\\_tamara@bls.gov](mailto:Zimmerman_tamara@bls.gov). <sup>c</sup> Department of Mathematics, Temple University, Philadelphia, PA 19122. E-mail: [mehta1007@comcast.net](mailto:mehta1007@comcast.net).

## Abstract

The precision of a small-domain sample estimator can be improved by estimating simultaneously its true value and sampling and non-sampling error components. In principle, this simultaneous estimation is superior to any two-step estimation of the true value and sampling error components, ignoring the non-sampling error component. In this paper, a time series model for state employment or unemployment is used to demonstrate the limitations of a two-step method. A cross-sectional model for state employment or unemployment is used to explain the advantages of simultaneously estimating a true value and the sums of sampling and non-sampling errors in two or more sample estimators of the true value.

**Key Words.** Indirect model-dependent estimate; domain indirect; time indirect; domain and time indirect; coefficient driver.

## 1. Introduction

Accurate estimates of employment and unemployment at various levels of geographic detail are needed to formulate good regional policies, such as the determination of eligibility for and/or the allocation of Federal resources, with a good understanding of local economic conditions. Such estimates cannot be produced from survey data collected in each small area because sample sizes within those areas are often either zero or too small to provide reliable estimates. If an area-specific sample is available but is not large enough to yield “direct estimates” of adequate accuracy, then their accuracy can be improved, using “indirect” model-dependent estimators that “borrow strength” from auxiliary data collected in this and other small areas and/or at more than one time period. On a definition suggested by a passage in Rao (2003, p. 2) indirect estimators might be classified as “domain indirect”, “time indirect”, or “domain and time indirect” depending on whether they borrow strength cross-sectionally, over time, or both. In this paper, we develop domain indirect estimators that facilitate simultaneous estimation of the true value and sampling and non-sampling error components of sample estimators. In principle, simultaneous estimation is superior to two-step estimation of these components. To illustrate the proposed techniques, we consider the problem of improving the accuracy of current population survey (CPS) estimates of employment and unemployment for 51 “small areas” consisting of the 50 United States plus the District of Columbia. CPS state estimates are widely used and improving their accuracy is valuable.

The remainder of this paper is divided into six sections. Section 2 describes the time series models used to estimate employment or unemployment for the states, major metropolitan areas, and corresponding balance of states in the U.S. Section 3 shows the limitations of a two-step method used to estimate these time series models. The consequences of incorrectly neglecting non-sampling errors are given in Section 4. Section 5 uses cross-sectional estimates based on two or more sample estimators of state employment (or unemployment) to estimate simultaneously the true values of employment (or unemployment) and the sums of sampling and non-sampling errors contained in those estimators. Seasonal adjustment of the true values of state employment (or unemployment) and estimation of the autocovariances of sampling errors are discussed in Section 6. Section 7 provides an example. Section 8 concludes.

## 2. Time Series Models for the True Value Component of a Sample Estimator

### 2.1 Sampling Model

---

**Acknowledgment.** Comments by Christopher Bollinger are highly appreciated.

**Disclaimer.** Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

Let  $\hat{Y}_{it}^{CPS}$  denote the current population survey (CPS) composite estimator of the true population value, denoted by  $Y_{it}$ , of either unemployment or employment.<sup>1</sup> Let  $i$  and  $j$  index states, major metropolitan areas, or balance of states and let  $t$  index months. The decomposition of  $\hat{Y}_{it}^{CPS}$  into its unobserved components is

$$\hat{Y}_{it}^{CPS} = Y_{it} + e_{it}^{CPS} + u_{it}^{CPS}, \quad t = 1, \dots, \tau_i, i = 1, \dots, m \quad (1)$$

where  $e_{it}^{CPS}$  and  $u_{it}^{CPS}$  represent sampling (or survey) and non-sampling errors (such as non-response or measurement errors), respectively. A value, denoted by  $y_{it}^{CPS}$ , of  $\hat{Y}_{it}^{CPS}$  obtained from a specific CPS sample is an estimate of  $Y_{it}$ .

The annual averages of the components of a type of non-response rates for 1993-1996 and 2003 CPS national estimates are given in U.S. Department of Labor (2006, p. 16-4). These averages show that we cannot assume that  $u_{it}^{CPS} = 0$ . The methods in Cochran (1977, Chapter 13) might have been already tried to reduce the magnitude of  $u_{it}^{CPS}$ . Under the current 2000 CPS design, the number of assigned households in monthly CPS samples for different states ranges from 700 to 5,344. These sample sizes are not large enough to yield  $\hat{Y}_{it}^{CPS}$  of adequate precision.<sup>2</sup> It is necessary to use model-based approaches to improve the efficiency of  $\hat{Y}_{it}^{CPS}$ .

The design-based approach to CPS inference treats the  $Y_{it}$  as fixed quantities. Unlike this approach, the model-based approach to CPS sampling inference treats the  $Y_{it}$  as a random sample from a “superpopulation” and assigns to them a probability distribution implied by a model for  $Y_{it}$ .<sup>3</sup> Since the “true” model for  $Y_{it}$  is unknown, we can use one of its approximations in place of  $Y_{it}$  used in (1) but then we should be prepared to carefully investigate whether such a use improves the precision of  $\hat{Y}_{it}^{CPS}$ .<sup>4</sup> Let  $E_p$  and  $V_p$  denote the expectation and variance operators with respect to the probability distribution induced by the 2000 CPS sampling design conditional on  $Y_{it}$ , respectively. Let  $E_m$  and  $V_m$  denote the expectation and variance operators with respect to a probability distribution assigned to the  $Y_{it}$  by an assumed model for  $Y_{it}$ , respectively (see Little, 2004, p. 547).

## 2.2 Survey Error (SurE) Model

Suppose that  $u_{it}^{CPS} = 0$  for all  $i$  and  $t$ . Then a model for  $e_{it}^{CPS}$  implied by the CPS design is

$$e_{it}^{CPS} = \frac{\sigma_{e^{CPS},it}}{\sigma_{e^*,i}} e_{it}^* \quad (2)$$

where for  $i, j = 1, \dots, m$  and  $t = 1, \dots, \tau_i$ ,  $E_p(e_{it}^{CPS} | Y_{it}) = 0$ ,  $Cov_p[(e_{it}^{CPS}, e_{jt}^{CPS}) | Y_{it}, Y_{jt}] = 0$  if  $i \neq j$ ,  $V_p(e_{it}^{CPS} | Y_{it}) = \sigma_{e^{CPS},it}^2 = \sigma_{b,e^{CPS},it}^2 + D_{e^{CPS},it} \sigma_{1e^{CPS},it}^2$  with  $\sigma_{b,e^{CPS},it}^2 =$  between-Primary Sampling Unit (PSU) variance component of the total design variance of  $e_{it}^{CPS}$  arising from selecting a subset of all non-self-representing (NSR) PSUs in state  $i$ ,  $D_{e^{CPS},it} \sigma_{1e^{CPS},it}^2 =$  within-PSU variance component of the total design variance of  $e_{it}^{CPS}$  arising from sampling of housing units within selected PSUs of state  $i$ ,  $D_{e^{CPS},it}$  is the ratio of the within-PSU variance component of the total design variance of  $e_{it}^{CPS}$ , assuming the CPS sample design, to the total design variance,  $\sigma_{1e^{CPS},it}^2 = (N_{it} / n_{it}) Y_{it} (1 - (Y_{it} / N_{it}))$ , of an unbiased estimator of  $Y_{it}$ , assuming simple random sampling (design effect),  $N_{it}$  is the civilian non-institutionalized population 16 years of age and older for the state,  $(N_{it} / n_{it})$  is the state sampling interval (see U.S. Department of Labor, 2006, p. 3-6) and  $e_{it}^*$  follows a mixed 2<sup>nd</sup> order

<sup>1</sup> Here employment is defined on a place-of-residence basis, since the CPS is a household survey.

<sup>2</sup> We strictly adhere to Cochran’s (1977, p. 16) distinction between “precision” and “accuracy”.

<sup>3</sup> Rao (2003, p. 77) calls this model a linking model.

<sup>4</sup> The “true” model for  $Y_{it}$  is an unknown function,  $Y_{it} = f_{it}(x_{1it}, \dots, x_{Kit}, x_{K+1,it}, \dots, x_{L_{it},it})$ , of all of the determinants of  $Y_{it}$  with the correct but unknown functional form and without any omitted determinants or mismeasured variables. Any of its estimable approximations is misspecified if it has an incorrect functional form and suffers from omitted-variable and measurement-error biases (see Freedman and Navidi, 1986).

autoregressive and 17<sup>th</sup> order moving average, or ARMA(2, 17), model with  $V_m(e_{it}^*) = \sigma_{e_{it}^*}^2$ , as shown by Tiller (1992, 2005). This ARMA(2, 17) model is approximated by a 15<sup>th</sup> order autoregression or AR(15) model.

### 2.3 Linking Model

Durbin and Koopman (2001) work with Harvey's (1989) structural time series models. Following them, it is assumed that the finite population is generated according to the superpopulation model

$$Y_{it} = T_{Y_{it}} + S_{Y_{it}} + I_{Y_{it}} \quad (3)$$

where  $T_{Y_{it}}$  is the trend-cycle,  $S_{Y_{it}}$  the seasonal, and  $I_{Y_{it}}$  the irregular component, which are treated as the unobserved components of  $Y_{it}$ .<sup>5</sup> All these components are assumed to be independent of each other for all  $i$  and  $t$ . This assumption may not be true if the decomposition of  $Y_{it}$  in equation (3) is not unique. To account for this non-uniqueness, Havenner and Swamy (1981) assume that  $T_{Y_{it}}$ ,  $S_{Y_{it}}$ , and  $I_{Y_{it}}$  are correlated with each other for all  $i$  and  $t$ .

It is assumed that  $T_{Y_{it}}$  follows a random walk model with one period lagged value of a random walk drift. Durbin and Koopman (2001, p. 39) call this the *local linear trend* model. The seasonal component,  $S_{Y_{it}}$ , is expressed in a trigonometric form to make it follow Durbin and Koopman's (2001, (3.6), p. 40) quasi-random walk model. The irregular component,  $I_{Y_{it}}$ , is assumed to be serially independent and normally distributed with mean zero and constant variance.<sup>6</sup>

It is known that employment has a strong tendency to move cyclically, downward in general business slowdowns and upward in expansions. These cycles are asymmetric because employment decreases at a faster rate than it increases. The behavior of unemployment over time is the opposite of employment's behavior and is called asymmetric counter-cyclical behavior. Montgomery, Zarnowitz, Tsay and Tiao's (1998, p. 487) results imply that the time series models assumed above for the components of  $Y_{it}$  in (3) may be misspecified because they may not exhibit the asymmetric cycles of employment and unemployment. More specifically, their results are: (i) A first order autoregressive model in first differences, denoted by ARIMA(1, 1, 0), that fitted the U.S. quarterly unemployment rate series for 1948-1993 quite well was not able to accurately represent the asymmetric cycles of unemployment during this period; (ii) An ARIMA (1, 1, 0)(4, 0, 4) model with a multiplicative seasonal factor, denoted by ARIMA(4, 0, 4), under-predicted the U.S. unemployment rate during the rapid increase of 1982 and exhibited forecasts that fluctuated a great deal more during stable periods of unemployment.<sup>7</sup>

### 2.4 Covariate Model

Consider an extension of model (3) to allow one or more of the unobserved components of  $Y_{it}$  to be related to corresponding components in another series called a covariate. The unemployment insurance claims (U.I. Claims) from the Federal-State Unemployment Insurance System are used as a covariate series if  $Y_{it}$  represents unemployment, and nonagricultural payroll employment estimates from the Current Employment Statistics (CES) survey are used as a covariate series if  $Y_{it}$  represents employment.

## 3. A Two-Step Method of Estimating the Parameters of Survey Error and Linking Models

We call the model consisting of equation (1) without  $u_{it}$ , equation (2), and equation (3) with covariates "the combined model". For each  $i$ , the unknown parameters of model (2) are: the design variances,  $\sigma_{e_{CPS, it}}^2$ ,  $t = 1, \dots, \tau_i$ , and 17 parameters of AR(15) model. For each  $i$ , the unknown parameters of model (3) are the error variances of the models for  $T_{Y_{it}}$ ,  $S_{Y_{it}}$ , and  $I_{Y_{it}}$ , which are four in number (see Tiller, 2005, p. 7). When the  $y_{it}^{CPS}$  for  $t = 1, \dots, \tau_i$ , are augmented with the available

<sup>5</sup> Tiller (2005) also allows for temporary or permanent shifts in the level of the series,  $\{Y_{it}\}$ ,  $t = -\infty, +\infty$ .

<sup>6</sup> Sometimes it is assumed that  $I_{Y_{it}}$  follows a lower-order autoregressive process (see Tiller, 1992, p. 151).

<sup>7</sup> It can be shown that  $Y_{it}$  contains a random walk component if it contains a stationary component after being first differenced  $d$  times. It is precisely this condition that is violated when the nonlinear models that accurately represent the asymmetric cycles of employment and unemployment are considered.

observations on the covariates, the number of the unknown parameters of (3) with covariates increases to 10 for each  $i$  (see Tiller, 2005, p. 10). All these parameters of the combined model are not identifiable on the basis of the available CPS estimates,  $y_{it}^{CPS}$ ,  $t = 1, \dots, \tau_i$ ,  $i = 1, \dots, m$ , and observations on the covariates alone. Consequently, for each  $i$ , a two-step procedure is used: First step. For each  $i$ , the available CPS estimates ( $y_{it}^{CPS}$ ,  $t = 1, \dots, \tau_i$ ) are used to estimate the parameters of (2) independently of (3) with covariates; Second step. Given the CPS estimates ( $y_{it}^{CPS}$ ,  $t = 1, \dots, \tau_i$ ), the parameters of (3) with covariates are estimated, holding the parameters of (2) and AR(15) fixed at their estimated values.

### 3.1 Survey Error Variance Estimates

Let  $i$  and  $t$  be fixed so that the components of  $Y_{it}$  in (3) are fixed. Let  $u_{it}^{CPS}$  be absent. Then the design variance,  $\sigma_{e^{CPS},it}^2$ , in (2) can be estimated using the generalized variance function (GVF) (see Wolter (1985, p. 203), Lent (1991, 1994), and U.S. Department of Labor (2006, p. 3-6)),

$$V_p(\hat{Y}_{it}^{CPS} | Y_{it}) = -(b_{it} / N_{it})Y_{it}^2 + b_{it}Y_{it} \quad (4)$$

where  $b_{it} = k_{it}(N_{it} / n_{it})D_{e^{CPS},it}$ ,  $k_{it} = 1/(1-p_{it})$ ,  $p_{it} = \sigma_{b,e^{CPS},it}^2 / \sigma_{e^{CPS},it}^2$ , and  $D_{e^{CPS},it}$ ,  $N_{it}$ ,  $n_{it}$ ,  $\sigma_{b,e^{CPS},it}^2$ , and  $\sigma_{e^{CPS},it}^2 (= V_p(\hat{Y}_{it}^{CPS} | Y_{it}))$  are as defined below (2).

A method of estimating the  $b_{it}$  is summarized in the following steps:

1. All survey statistics are divided into several groups with model (4) fitted independently in each group. If  $b_{it}$  differs among the statistics in a group, then they are not consistently estimable because the number of unknown parameters increases with the size of the group (see Lehmann and Casella, 1998, p. 481). To avoid this difficult *incidental parameter* problem, care is taken to group together all survey statistics that follow a common model such as (4) with fixed  $i$  and  $t$ . This may involve grouping together the statistics with similar design effects; the same characteristics for selected demographic or geographic subgroups. Examples of such groupings are given in Wolter (1985, p. 209).
2. Let  $g$  index the statistics in a group formed in Step 1. The design variances,  $V_p(\hat{Y}_{it}^{CPS}(g) | Y_{it}(g))$ , are computed for several statistics of this group using  $-(\hat{b}_{it} / N_{it})(\hat{Y}_{it}^{CPS}(g))^2 + \hat{b}_{it}\hat{Y}_{it}^{CPS}(g)$ , where  $\hat{b}_{it} = \hat{k}_{it}(N_{it} / n_{it})\hat{D}_{e^{CPS},it}$ ,  $\hat{k}_{it} = 1/(1-\hat{p}_{it})$ ,  $\hat{p}_{it} = \hat{\sigma}_{b,e^{CPS},it}^2 / \hat{\sigma}_{e^{CPS},it}^2$ ,  $\hat{\sigma}_{e^{CPS},it}^2 = \hat{\sigma}_{b,e^{CPS},it}^2 + \hat{D}_{e^{CPS},it}\hat{\sigma}_{1e^{CPS},it}^2$ ,  $\hat{\sigma}_{b,e^{CPS},it}^2$  is obtained by multiplying the 2000 decennial census estimate of between-PSU variance by the square of the ratio of the annual average of  $\hat{Y}_{it}^{CPS}(g)$  for the current year to the 2000 census estimate of  $Y_{it}(g)$ ,  $\hat{D}_{e^{CPS},it}$  is computed from the national within-PSU design effects by adjusting for differences in a certain noninterview rate, see the memorandum (2007) from Khandaker Mansur, and  $\hat{\sigma}_{1e^{CPS},it}^2 = \sigma_{1e^{CPS},it}^2$  when  $Y_{it}(g) = \hat{Y}_{it}^{CPS}(g)$ . Let  $\hat{V}_p(\hat{Y}_{it}^{CPS}(g) | Y_{it}(g))$  denote the estimator of  $V_p(\hat{Y}_{it}^{CPS}(g) | Y_{it}(g))$ .
3. Fitting model (4) to the data  $(\hat{Y}_{it}^{CPS}(g), \hat{V}_p(\hat{Y}_{it}^{CPS}(g) | Y_{it}(g)))$  from Step 2 gives the estimate of  $b_{it}$  for the group. The model fitting technique is an iterative weighted least squares procedure, where the weight is the inverse of the square of the predicted value of  $\hat{V}_p(\hat{Y}_{it}^{CPS}(g) | Y_{it}(g)) / (\hat{Y}_{it}^{CPS}(g))^2$  (see Wolter, 1985, p. 207, (5.4.2)). Let  $\tilde{b}_{it}$  denote an iterative weighted least squares estimator of  $b_{it}$ .
4. An estimate of the design variance of a survey statistic, say  $\hat{Y}_{it}^{CPS}(g)$ , for which the successive difference replication method of Fay and Train (1995) is not applied is now obtained by evaluating model (4) at the point  $(\hat{Y}_{it}^{CPS}(g); \tilde{b}_{it})$ . It is called a GVF estimate.

The estimator,  $\tilde{b}_{it}$ , obtained in Step 3 may not have any desirable statistical properties because of the effects of the errors,  $\hat{Y}_{it}^{CPS}(g) - Y_{it}(g)$ , and can be very imprecise if all statistics within a group do not behave according to model (4). These problems may get resolved if model (4) is replaced by the less restrictive model,

$$\hat{V}_p(\hat{Y}_{it}^{CPS}(g) | Y_{it}(g)) = b_{itg} \{ \hat{Y}_{it}^{CPS}(g) - (1/N_{it})(\hat{Y}_{it}^{CPS}(g))^2 \}, \quad g = 1, 2, \dots, G \quad (5)$$

with

$$b_{itg} = \delta_{01}\hat{k}_{itg}(N_{it} / n_{it})\hat{D}_{e^{CPS},itg} + \xi_{0itg} \quad (6)$$

where  $\hat{D}_{e^{CPS},itg}$  and  $\hat{k}_{itg}$  are the estimators of  $D_{e^{CPS},itg}$  and  $k_{itg}$  given in Step 2, and  $\xi_{0itg}$  is a random variable with mean zero and constant variance. Chang, Swamy, Hallahan and Tavlas' (2000) method can be used to estimate model (5) under assumption (6). Estimator (5) is imprecise, since  $\hat{Y}_{it}^{CPS}(g)$  is an imprecise estimator of  $Y_{it}(g)$ .

### 3.2 Survey Error Autocorrelation Estimates

In Zimmerman and Robison (1995), the variances and autocovariances of the  $e_{it}^*$  in (2) are estimated from the separate panel (rotation group) estimates of  $Y_{it}$ . A panel is defined as the set of sampling units joining and leaving the sample at the same time. Each panel is a representative sample of the population. The CPS sample consists of 8 such panels in every month. To estimate model variance and autocovariances of the CPS composite estimator from the estimates of model variances and autocovariances of the panel estimators, Zimmerman and Robison (1995) and Zimmerman (2007) consider the following model:

$$\hat{Y}_{ijt}^{CPS} = \mu_i + \theta_{ij} + \beta_{it} + e_{ijt}^{CPS} \quad (7)$$

where  $\hat{Y}_{ijt}^{CPS}$  = CPS estimator of  $Y_{it}$  from panel  $j$ ,  $\mu_i$  = overall mean,  $\theta_{ij}$  = month-in-sample effect,  $\beta_{it}$  = time effect, and  $e_{ijt}^{CPS}$  = sampling error of  $\hat{Y}_{ijt}^{CPS}$ .

Let  $e_{ijt}^{CPS}$  denote  $e_{it}^{CPS}$  in (2) for panel  $j$ . Model (7) is misspecified if the time series models for the components of  $Y_{it}$  in (3) are correctly specified. If this is so, model (7) incorrectly estimates  $e_{ijt}^{CPS}$  as  $\hat{Y}_{ijt}^{CPS} - \hat{\mu}_i - \hat{\theta}_{ij} - \hat{\beta}_{it}$  instead of as  $\hat{Y}_{ijt}^{CPS} - \hat{T}_{y_{ijt},ijt} - \hat{S}_{y_{ijt},ijt} - \hat{I}_{y_{ijt},ijt}$ . The incorrectly estimated  $e_{ijt}^{CPS}$  cannot give *accurate* estimates of the parameters of model (2) other than  $\sigma_{e^{CPS},it}^2$  even if model (5) gives accurate estimates of  $\sigma_{e^{CPS},it}^2$ . Thus, the First step--of estimating the survey error model parameters from design based information independently of the time series models for the components of  $Y_{it}$  in (3)--can lead to very inaccurate estimates of the survey error model parameters. If the condition, "holding the parameters of the survey error model fixed at their estimated values", means the condition, "the parameters of the survey error model are set equal to the estimates obtained in the First step", then the inaccurate estimates obtained in the First step can lead to inaccurate estimates of the parameters of model (3) with covariates in the Second step. There will be no consistency between the estimates obtained in the First step and those obtained in the Second step. Furthermore, the random walk models assumed for some of the components of  $Y_{it}$  lead to a predictor of  $Y_{it}$  that is *unconditionally inadmissible* relative to quadratic loss functions because the predictor does not possess finite unconditional mean. They do not provide the predictors of  $Y_{it}$  with good *conditional* and *unconditional* properties. Brown (1990, p. 491) shows the importance of working with such predictors.

### 3.3 State-Space Form of the Combined Model

Let  $\{x_{it}\}$  be a covariate series. Let  $y_{xit} = (y_{it}^{CPS}, x_{it})'$ . Let the combined model be written in state-space form, as in Durbin and Koopman (2001, p. 38). Hold the parameters of the survey error model fixed at the estimates obtained in the First step. For calculating in the Second step loglikelihood function and the maximization of it with respect to the parameters of model (3) with covariates, the joint density of the sample observations,  $y_{xi1}, y_{xi2}, \dots, y_{xi\tau_i}$ , implied by the combined model is written as

$$p(y_{xi1}, \dots, y_{xi\tau_i}) = \prod_{t=1}^{\tau_i} p(y_{xit} | y_{xi0}, \dots, y_{xi1}) \quad (8)$$

where it is assumed that  $p(y_{xi1} | y_{xi0}) = p(y_{xi1})$  with finite mean (see Durbin and Koopman, 2001, p. 30). The truth of this assumption is usually unknown. For this reason, it is convenient to assume that  $\tilde{Y}_{xit}$  is distributed with density  $p(y_{xit} | y_{xi0}, \dots, y_{xi1})$  for all  $t$ ,<sup>8</sup> unfortunately, this assumption contradicts the assumption that  $p(y_{xi1} | y_{xi0}) = p(y_{xi1})$  with finite mean. This is because the former assumption says that under the random walk models assumed for some of the components of  $Y_{it}$ ,  $\tilde{Y}_{xit}$  does not possess finite *unconditional mean* for all  $t$  and the latter assumption says that  $\tilde{Y}_{xit}$  possesses

<sup>8</sup> Here we distinguish a random variable from its value by a tilde. For example,  $y_{xit}$  is the value taken by the random variable  $\tilde{Y}_{xit} = (\hat{Y}_{it}^{CPS}, X_{it})'$ .

finite *unconditional mean* for  $t = 1$ . Even if it is assumed that  $\tilde{Y}_{xit}$  is distributed with density  $p(y_{xit} | y_{xit-1}, \dots, y_{xit1})$  for all  $t > 1$ , the assumption that  $p(y_{xit1} | y_{xit0}) = p(y_{xit1})$  with finite mean cannot be true for all those data sets for which the value  $t = 1$  occurs at different points on the time axis.

One may use  $l (> 0)$  non-stationary elements in the state vector which determines the number of observations required to form priors of these elements. Without this or any other initialization, it is not possible to apply the Kalman filter to state-space form of the combined model in (1)-(3) (see Durbin and Koopman (2001, pp. 27-30, 99-104) and Maddala and Kim (1998, pp. 475-477)). The initial value  $y_{xit0}$  is usually unknown and any assumption about it may be questioned. Incorrect assumptions about the  $y_{xit0}$  can lead to the estimates of  $Y_{it}$ ,  $t = 1, \dots, \tau_i$ , with incorrect initial values and hence with incorrect time profiles.

#### 4. Consequences of Incorrectly Neglecting Non-Sampling Errors

In the previous section, we have considered equation (1) with  $u_{it}^{CPS}$  suppressed even though  $u_{it}^{CPS}$  is not equal to zero with probability 1. Adding to this equation a linking model of the general linear mixed (GLM) model's type for  $Y_{it}$ , which implies that  $Y_{it}$  possesses finite second moment for all  $i$  and  $t$ , the best linear unbiased predictor (BLUP) of  $Y_{it}$  can be found.<sup>9</sup> This BLUP can be expressed as a weighted average of  $\hat{Y}_{it}^{CPS}$  and the regression-synthetic estimator of  $Y_{it}$  implied by the linking model. Under certain regularity conditions given in Rao (2003, p. 117), the BLUP of  $Y_{it}$  coincides with  $\hat{Y}_{it}^{CPS}$  that is not affected by the misspecifications in the linking model as the design variance of  $\hat{Y}_{it}^{CPS}$  goes to 0. That is, when  $\hat{Y}_{it}^{CPS}$  is design-consistent, the BLUP of  $Y_{it}$  can also be design-consistent. This result is the basis of Little's (2004, p. 551) statement that one way of limiting the effects of model misspecification is to restrict attention to models that yield design-consistent estimators (see also Rao, 2003, p. 148). This observation is of no use to us when  $\hat{Y}_{it}^{CPS}$  is subject to non-sampling errors, in which case  $\hat{Y}_{it}^{CPS}$  is not design consistent (see Little, 2004, p. 549).

In Rao's (2003) terminology, the BLUP of  $Y_{it}$  becomes an empirical BLUP (EBLUP) when the variance parameters involved in the BLUP are replaced by their respective sample estimators. A second-order approximation to the mean square error (MSE) of EBLUP involves three terms, as Rao's (2003, p. 104, (6.2.31)) very elegant derivation shows.<sup>10</sup> This derivation further shows that an unbiased estimator of the MSE of EBLUP to a desired order of approximation may involve four terms (see Rao, 2003, p. 105, (6.2.37)). The value of the sum of these four terms obtained from a specific sample can exceed the value of the design variance of  $\hat{Y}_{it}^{CPS}$ , showing that the EBLUP of  $Y_{it}$  can be less efficient than  $\hat{Y}_{it}^{CPS}$  which is itself imprecise because of the smallness of the sample on which it is based. The MSE of the BLUP of  $Y_{it}$  contains two terms and only one of these terms is smaller than the design variance of  $\hat{Y}_{it}^{CPS}$ , as shown by Rao (2003, p. 117, (7.1.7)). These results raise the question: in what sense does the linking model for  $Y_{it}$  "borrow strength" from its explanatory variables in making an estimate of  $Y_{it}$ ? Under certain regularity conditions, two of the three terms of a second-order approximation to the MSE of an EBLUP of  $Y_{it}$  go to zero and the remaining term remains below the design variance of  $\hat{Y}_{it}^{CPS}$  as the number of observations on the explanatory variables of the linking model goes to infinity (see Rao, 2003, p. 117). Thus, in the limit the MSE of an EBLUP of  $Y_{it}$  can involve only one term that is smaller than the design variance of  $\hat{Y}_{it}^{CPS}$ .<sup>11</sup> Any linking model for  $Y_{it}$  can borrow strength from its explanatory variables in the limit as the number of observations on its explanatory variables goes to infinity if  $\hat{Y}_{it}^{CPS}$  is design-consistent and some regularity conditions are satisfied. This result holds even when the linking model for

<sup>9</sup> Swamy, Zimmerman, Mehta and Robison (2005) extend Rao's (2003, pp. 96-98) derivation of the BLUP from a GLM model for a sample estimator to take account of omitted-variable, measurement-error and incorrect functional-form (or simply specification) biases in the GLM model.

<sup>10</sup> This derivation has been extended to take account of specification biases in linking models of the GLM type in Swamy, Yaghi, Mehta and Chang (2006).

<sup>11</sup> The gain in efficiency associated with the use of Rao's (2003, pp. 116-117) BLUP may get reduced if the linking model that provides the BLUP suffers from specification biases (see Swamy et al., 2005).

$Y_{it}$  is misspecified, provided  $\hat{Y}_{it}^{CPS}$  does not contain non-sampling errors. It is of limited use for us because large sample theories of estimation are irrelevant to small area estimation and the sources of error other than random sampling variation given in Cochran (1977, p. 359) are present in the CPS and CES survey. There is usually no justification to ignore the possible misspecifications in the time series models for the components of  $Y_{it}$  in equation (3) with covariates.

## 5. Simultaneous Estimation of Sums of Survey and Non-Sampling Errors and the Corresponding True Values

Let us now reconsider  $\hat{Y}_{it}^{CPS}$ . Suppose that an additional estimator, denoted by  $\hat{Y}_{it}^A$ , of  $Y_{it}$  is available. This additional estimator will be given by U.I. Claims data if  $Y_{it}$  represents unemployment and by data from the CES survey or by extrapolated data from the Quarterly Census of Employment and Wages (QCEW) program if  $Y_{it}$  represents employment. We assume that each of  $\hat{Y}_{it}^{CPS}$  and  $\hat{Y}_{it}^A$  is subject to both sampling and non-sampling errors and write  $\hat{Y}_{it}^{CPS} = Y_{it} + \varepsilon_{it}^{CPS}$  with  $\varepsilon_{it}^{CPS} = e_{it}^{CPS} + u_{it}^{CPS}$  and  $\hat{Y}_{it}^A = Y_{it} + \varepsilon_{it}^A$  with  $\varepsilon_{it}^A = e_{it}^A + u_{it}^A$ , where the vectors,  $(e_{it}^{CPS}, e_{it}^A)'$  and  $(u_{it}^{CPS}, u_{it}^A)'$ , denote the sampling and non-sampling errors of  $(\hat{Y}_{it}^{CPS}, \hat{Y}_{it}^A)'$ , respectively.<sup>12</sup> Let  $y_{it}^A$  denote a value of  $\hat{Y}_{it}^A$  obtained from a specific sample.

We now fix  $t$  and let  $i$  vary to focus on cross-sectional variations in  $\varepsilon_{it}^{CPS}$  and  $\varepsilon_{it}^A$ . These types of variations in  $e_{it}^{CPS}$  and  $e_{it}^A$  are different from their random sampling variations that are present because only parts of the population have been measured using some sampling designs.<sup>13</sup> Modeling assumptions are needed to analyze the non-sampling errors,  $u_{it}^{CPS}$  and  $u_{it}^A$ , and cross-sectional variations in the sampling errors,  $e_{it}^{CPS}$  and  $e_{it}^A$ . The problem of choosing between  $\hat{Y}_{it}^{CPS}$  and  $\hat{Y}_{it}^A$  can only be solved if we can derive from them a single estimate that is closer to  $Y_{it}$  than either estimate. We show in this section that such a single estimate can be found even when  $|y_{it}^A - Y_{it}| > |y_{it}^{CPS} - Y_{it}|$ .

### 5.1 Random Coefficient Regression Model

Writing  $\alpha_{1it}^A = 1 - (\varepsilon_{it}^A / \hat{Y}_{it}^A)$ , a function of  $\hat{Y}_{it}^A$  and  $\varepsilon_{it}^A$ , we transform the sampling model,  $\hat{Y}_{it}^A = Y_{it} + \varepsilon_{it}^A$ , into the linking model,  $Y_{it} = \alpha_{1it}^A \hat{Y}_{it}^A$ . Replacing  $Y_{it}$  in the sampling model in (1) by  $\alpha_{1it}^A \hat{Y}_{it}^A$  gives the model:

$$\hat{Y}_{it}^{CPS} = \alpha_{0it}^{CPS} + \alpha_{1it}^A \hat{Y}_{it}^A, \quad i = 1, \dots, m \quad (9)$$

where  $\alpha_{0it}^{CPS} = \varepsilon_{it}^{CPS}$ ,

$$\alpha_{0it}^{CPS} = \pi_{00}^{ca} + \pi_{01}^{ca} BP_{it} + \pi_{02}^{ca} HP_{it} + \zeta_{0it}^{ca} \quad (10)$$

$$\alpha_{1it}^A = \pi_{10}^{ca} + \pi_{11}^{ca} \left( \frac{BP_{it}}{TP_{it}} \right) + \pi_{12}^{ca} \left( \frac{HP_{it}}{TP_{it}} \right) + \zeta_{1it}^{ca} \quad (11)$$

where  $BP$  = the black population 16 years of age and older for the area,  $HP$  = the Hispanic population 16 years of age and older for the area,  $TP$  = the civilian non-institutionalized population 16 years of age and older for the area, all the  $\pi$  s are fixed, the superscript,  $ca$ , of the  $\pi$  s and  $\zeta$  s is shorthand for “regression of  $\hat{Y}_{it}^{CPS}$  on  $\hat{Y}_{it}^A$ ”, and the variables,  $BP$ ,  $HP$ ,  $\left( \frac{BP}{TP} \right)$ ,

<sup>12</sup> Note that if  $\hat{Y}_{it}^A$  is an estimator of employment, then its correct decomposition is  $\hat{Y}_{it}^A = Y_{it} + (Y_{it}^* - Y_{it}) + e_{it}^A + u_{it}^A$ , where  $Y_{it}^*$  is a “place-of-work” employment that is different from the place-of-residence employment,  $Y_{it}$ , the true-value component of  $\hat{Y}_{it}^{CPS}$ . To adopt the “place-of-residence” concept, the difference,  $Y_{it}^* - Y_{it}$ , is added to  $u_{it}^A$  so that  $\hat{Y}_{it}^A = Y_{it} + \varepsilon_{it}^A$ . Mathiowetz and Duncan (1988) adopt the “place-of-work” concept to study response errors in retrospective reports of unemployment.

<sup>13</sup> Here we ignore the autocorrelations of  $e_{it}^{CPS}$  implied by the CPS sampling design because with fixed  $t$  and varying  $i$ ,  $e_{it}^{CPS}$  exhibits only cross-sectional variation. Our purpose in this section is to model this variation in the presence of non-sampling errors. An advantage of having cross-sectional data alone on two or more estimators of  $Y_{it}$  is that they can provide information about the sum of the sampling and non-sampling errors of  $\hat{Y}_{it}^{CPS}$ , whereas time-series data alone on  $\hat{Y}_{it}^{CPS}$  muddle the two errors, with no prospect of estimating even their sum. The former result follows from the analysis of this section and the latter result follows from the analysis of Section 3.

and  $\left(\frac{HP}{TP}\right)$ , are called “the coefficient drivers”.<sup>14</sup> Additional coefficient drivers may be included in (10) and (11) if they are thought to be appropriate.

We estimate equations (9)-(11) under the following assumptions: For  $i = 1, \dots, m$  and fixed  $t$ ,

(A1) the  $(\zeta_{0it}^{ca}, \zeta_{1it}^{ca})'$  are conditionally and independently distributed with mean vector 0 and constant covariance matrix

$\sigma_\zeta^2 \Delta_\zeta$  given the coefficient drivers.

(A2)  $P(\alpha_{0it}^{CPS}, \alpha_{1it}^A | \hat{Y}_{it}^A, BP_{it}, HP_{it}, TP_{it}) = P(\alpha_{0it}^{CPS}, \alpha_{1it}^A | BP_{it}, HP_{it}, TP_{it})$  where  $P(\cdot)$  is a joint probability distribution function of the variables in (9)-(11).

The costs of assumptions (A1) and (A2) are considerably less than those of assumptions underlying the combined model in Section 3, as the following discussion shows:

- (i) Design-based inference is strictly inapplicable to situations where  $u_{it}^{CPS} \neq 0$  (see Little, 2004, p. 540). Hence model (2) that is needed for design-based inference is not employed, since such situations are considered in this section.
- (ii) What we have done in (9) is that we retained the sampling model (1) and avoided all the misspecifications in, and the identification problems with, models (2) and (3) by replacing the linking model (3) by its alternative,  $Y_{it} = \alpha_{1it}^A \hat{Y}_{it}^A$ . Not much is known about the true model for  $Y_{it}$  in footnote 4. Therefore, the correctness of any specified model for  $Y_{it}$  may be questioned. However, the misspecifications in the time series models for the components of  $Y_{it}$  in (3) seemed to us to cry out for alternative linking models. A careful comparison of the results produced by model (3) and its alternative,  $Y_{it} = \alpha_{1it}^A \hat{Y}_{it}^A$ , might expose the weaknesses of both models.
- (iii) The conditional independence assumption (A2) is weaker than the unconditional independence assumption,  $P(\alpha_{0it}^{CPS}, \alpha_{1it}^A | \hat{Y}_{it}^A) = P(\alpha_{0it}^{CPS}, \alpha_{1it}^A)$ , which is false because of the nonzero correlations between  $(\hat{Y}_{it}^A, e_{it}^A, u_{it}^A)'$  and  $\alpha_{1it}^A$ . We include the coefficient drivers in (10) and (11) to avoid this false assumption.
- (iv) Since (9) is a relation between two estimators of employment or unemployment, the range of its intercept is much wider than that of its slope. To account for this difference in ranges, its intercept is made to depend on  $BP_{it}$  and  $HP_{it}$ , which take values over wider intervals than  $\frac{BP_{it}}{TP_{it}}$  and  $\frac{HP_{it}}{TP_{it}}$ , on which its slope is made to depend.
- (v) Simultaneous estimation of the intercept and slope of (9) permits simultaneous estimation of all the components of  $\hat{Y}_{it}^{CPS}$  in (1), since  $\alpha_{0it}^{CPS} = \varepsilon_{it}^{CPS}$  and  $Y_{it} = \alpha_{1it}^A \hat{Y}_{it}^A$ . In principle, this simultaneous estimation is superior to the two-step estimation of  $Y_{it}$  and  $e_{it}^{CPS}$ , ignoring  $u_{it}^{CPS}$ , as in Section 3. The simultaneous estimation based on (9)-(11) eliminates the inconsistencies inherent in the two-step estimation of Section 3. The Chebychev inequality shows that if the variance of  $\zeta_{0it}^{ca}$  is small, then the distribution of  $\alpha_{0it}^{CPS}$  is tightly concentrated about the right-hand side of (10) with the error suppressed (see Lehmann, 1999, p. 52). In this case, the right-hand side of (10) with the error suppressed gives a good determination of the sum of sampling and non-sampling errors in  $\hat{Y}_{it}^{CPS}$ . Subtracting a good estimator of  $e_{it}^{CPS} + u_{it}^{CPS}$  from  $\hat{Y}_{it}^{CPS}$  gives a good estimator of  $Y_{it}$ .<sup>15</sup> Thus, under a restriction on the variance of  $\zeta_{0it}^{ca}$ , equations (9)-(11) permit accurate simultaneous estimation of all the components of  $\hat{Y}_{it}^{CPS}$ . These results reveal the advantages of using equations (9)-(11) instead of the combined model in Section 3.
- (vi) Swamy, Mehta and Chang (2006) show that superior estimates of  $Y_{it}$  are unlikely to be obtained when the dependent and explanatory variables of (9) are interchanged. Therefore, we do not make such an interchange in this paper.

For  $i = 1, \dots, m$  and fixed  $t$ , the  $Y_{it}$  are assumed to be a random sample from the “superpopulation” defined by  $Y_{it} = \alpha_{1it}^A \hat{Y}_{it}^A$  and are assigned a model conditional probability distribution of  $(\hat{Y}_{it}^{CPS} - \alpha_{0it}^{CPS}) = \alpha_{1it}^A \hat{Y}_{it}^A$  given  $\hat{Y}_{it}^A = y_{it}^A$  and the coefficient

<sup>14</sup> Swamy, Mehta and Chang (2006) extend model (9) to include more than two estimators of  $Y_{it}$  and also to sub-state areas for which CPS data are either too sparse or unavailable.

<sup>15</sup> The importance of having a good estimate of employment can be seen from Rowthorn and Glyn ((2006) and Magnac and Visser (1999).

drivers. Assumptions (10) and (11) imply that  $\varepsilon_{it}^{CPS}$  ( $= \alpha_{0it}^{CPS}$ ) and  $\varepsilon_{it}^A = \hat{Y}_{it}^A - Y_{it} = (1 - \alpha_{1it}^A) \hat{Y}_{it}^A$  follow a bivariate conditional frequency distributions given  $\hat{Y}_{it}^A = y_{it}^A$  and the coefficient drivers, as  $i$  varies for fixed  $t$ . Let the model means of this distribution be denoted by  $E_m(\varepsilon_{it}^{CPS} | \hat{Y}_{it}^A = y_{it}^A, BP_{it}, HP_{it}) = \mu_{it}^{CPS}$  and  $E_m(\varepsilon_{it}^A | \hat{Y}_{it}^A = y_{it}^A, \frac{BP_{it}}{TP_{it}}, \frac{HP_{it}}{TP_{it}}) = \mu_{it}^A$ . Then  $\mu_{it}^{CPS}$  and  $\mu_{it}^A$  represent the biases in  $\hat{Y}_{it}^{CPS}$  and  $\hat{Y}_{it}^A$ , respectively. The connection between  $(\mu_{it}^{CPS}, \mu_{it}^A)'$  and the terms in equations (10) and (11) is as follows:

$$\begin{aligned} \pi_{00}^{ca} &= \text{Constant bias in } \hat{Y}_{it}^{CPS} \text{ that affects all } i \text{ or the constant term of the bias, } \mu_{it}^{CPS}, \text{ in } \hat{Y}_{it}^{CPS}; \\ \pi_{01}^{ca} BP_{it} + \pi_{02}^{ca} HP_{it} &= \text{Variable component of the bias, } \mu_{it}^{CPS}, \text{ in } \hat{Y}_{it}^{CPS}; \\ \zeta_{0it}^{ca} &= \text{Fluctuating component of the error, } \varepsilon_{it}^{CPS}; \\ [(1 - \pi_{10}^{ca}) - \pi_{11}^{ca} (\frac{BP_{it}}{TP_{it}}) - \pi_{12}^{ca} (\frac{HP_{it}}{TP_{it}})] y_{it}^A &= \text{Variable component of the bias, } \mu_{it}^A, \text{ in } \hat{Y}_{it}^A; \\ -\zeta_{1it}^{ca} y_{it}^A &= \text{Fluctuating component of the error, } \varepsilon_{it}^A. \end{aligned}$$

Thus,  $\alpha_{1it}^A$  presents no difficult problems even though it is a non-linear function of  $\hat{Y}_{it}^A$  and  $\varepsilon_{it}^A$  if assumption (A2) is true. The above decomposition of the sampling and non-sampling errors of  $\hat{Y}_{it}^{CPS}$  is related to Cochran's (1977, p. 378) decomposition of an error of measurement on a unit based on his model of measurement error.<sup>16</sup> It should be noted that with fixed  $t$ , the time series shocks of  $\hat{Y}_{it}^{CPS}$  and  $\hat{Y}_{it}^A$  are fixed and get subsumed into  $\mu_{it}^{CPS}$  and  $\mu_{it}^A$ , respectively.

Inserting (10) and (11) into model (9) gives

$$\begin{aligned} \hat{Y}_{it}^{CPS} &= (\pi_{00}^{ca} + \pi_{01}^{ca} BP_{it} + \pi_{02}^{ca} HP_{it} + \zeta_{0it}^{ca}) \\ &+ \{ \pi_{10}^{ca} + \pi_{11}^{ca} (\frac{BP_{it}}{TP_{it}}) + \pi_{12}^{ca} (\frac{HP_{it}}{TP_{it}}) + \zeta_{1it}^{ca} \} \hat{Y}_{it}^A, \quad i = 1, \dots, m. \end{aligned} \quad (12)$$

Note that in this model, the interactions between some of the coefficient drivers and the explanatory variable of model (9) appear as additional explanatory variables and the disturbances,  $\zeta_{0it}^{ca} + \zeta_{1it}^{ca} \hat{Y}_{it}^A$ , are heteroscedastic. This means that the formulation in (12) is much richer than model (9) with fixed intercept and fixed slope which is misspecified.<sup>17</sup> In (12), cross-sectional variation in the pair,  $(\hat{Y}_{it}^{CPS}, \hat{Y}_{it}^A)$ , is modeled by getting the national intercept,  $\pi_{00}^{ca}$ , and slope,  $\pi_{10}^{ca}$ ,--which are common to all areas--modified by individual area components. In other words, the  $m$  pairs,  $(\hat{Y}_{it}^{CPS}, \hat{Y}_{it}^A)$ , are modeled simultaneously so that model (12) for each area consists of the common national intercept,  $\pi_{00}^{ca}$ , and slope,  $\pi_{10}^{ca}$ , and their deviations from the national level. Model (12) recognizes major aspects of the sampling designs yielding  $\hat{Y}_{it}^{CPS}$  and  $\hat{Y}_{it}^A$  because the coefficient drivers are selected to explain high proportions of spatial and temporal variations in  $Y_{it}$  and the sums of sampling and non-sampling errors of  $\hat{Y}_{it}^{CPS}$  and  $\hat{Y}_{it}^A$ . On a definition suggested by a sentence in Hwang and Dempster (1999, p. 298) attempting to recognize major aspects of sample design, temporal variation, and spatial variation amounts to getting the science right. Model (12) gets the science right in this sense. The misspecifications in (12) with these features can be less serious than those in (3).

Under assumptions (10) and (11), model (12) implies that for  $i = 1, \dots, m$  and fixed  $t$ , given  $\hat{Y}_{it}^A = y_{it}^A$  and the coefficient drivers, the  $\hat{Y}_{it}^{CPS}$  are independently and conditionally distributed with model mean equal to  $\mu_{it}^{CPS} - \mu_{it}^A + y_{it}^A$  and model variance equal to  $V_m\{(\zeta_{0it}^{ca} + \zeta_{1it}^{ca} \hat{Y}_{it}^A) | \hat{Y}_{it}^A = y_{it}^A, BP_{it}, HP_{it}, TP_{it}\}$ . The basic fitting algorithm for (12) is an Iteratively Re-Scaled Generalized Least Squares (IRSGLS) method of Chang, et al. (2000) where in every iteration, the weighting of the sums of squares and cross products of observations on the dependent and explanatory variables in (12) by the elements of the inverse

<sup>16</sup> Estimation of this model requires replicated data which we do not have.

<sup>17</sup> The treatment of the coefficients of model (9) as constants is a misspecification because it ignores (i) variations in the sampling and non-sampling errors of  $\hat{Y}_{it}^A$ , (ii) variations in the sampling and non-sampling errors of  $\hat{Y}_{it}^{CPS}$ , and (iii) the correlations between  $\alpha_{1it}^A$  and the elements of  $(\hat{Y}_{it}^A, e_{it}^A, u_{it}^A)'$ .

of the covariance matrix of the heteroscedastic disturbances,  $\zeta_{0it}^{ca} + \zeta_{1it}^{ca} y_{it}^A$ , will be performed. These model-based weights as well as the design-based weights embedded in  $\hat{Y}_{it}^{CPS}$  and  $\hat{Y}_{it}^A$  affect the estimates of the coefficients and the error terms of (12). This is how the sampling mechanisms generating  $\hat{Y}_{it}^{CPS}$  and  $\hat{Y}_{it}^A$  are modeled in (12). A model-based approach that ignores the sampling mechanism is not valid unless the sampling distribution does not depend on the survey outcomes (see Little, 2004, p. 548).

## 5.2 Estimation of Model (12)

The number of the unknown parameters of model (12) is 10: six  $\pi$ s, three distinct elements of  $\Delta_\zeta$ , and one  $\sigma_\zeta^2$ . Given the cross-sectional data,  $y_{it}^{CPS}$ ,  $y_{it}^A$ ,  $BP_{it}$ ,  $HP_{it}$ ,  $TP_{it}$ , for  $i = 1, \dots, m$  and fixed  $t$ , an IRSGLS method can be used to obtain good approximations to the minimum variance linear unbiased estimators (MVLUEs) of the coefficients and the BLUPs of the errors of model (12). These approximations are denoted by  $(\hat{\pi}_{kh}^{ca}, k = 0, 1, h = 0, 1, 2, \hat{\zeta}_{kit}^{ca}, k = 0, 1)$ . For  $k = 0, 1$ ,  $\hat{\zeta}_{kit}^{ca}$  is an EBLUP of  $\zeta_{kit}^{ca}$ . Using these approximations in place of their population counterparts used in (10) and (11) gives the estimators of  $\alpha_{0it}^{CPS}$  and  $\alpha_{1it}^A$ , denoted by  $\hat{\alpha}_{0it}^{CPS}$  and  $\hat{\alpha}_{1it}^A$ , respectively. The corresponding estimator of  $\sigma_\zeta^2 \Delta_\zeta$  is denoted by  $\hat{\sigma}_\zeta^2 \hat{\Delta}_\zeta$ . Appropriate formulas for computing the standard errors of these estimates have been worked out in the Appendix.

Kariya and Kurata (2004, pp. 42 and 73) prove that the IRSGLS estimators of the coefficients and the EBLUPs of the errors of model (12) possess finite second-order moments under very general conditions. To these conditions the condition that  $m - 6 \geq 6$  should be added. With this additional condition, two or more degrees of freedom remain unutilized after the estimation of the coefficients of (12),  $\sigma_\zeta^2$ , and the elements of  $\Delta_\zeta$ . Sufficient conditions for the consistency and asymptotic normality of the IRSGLS estimators of the coefficients of model (12) are given in Cavanagh and Rothenberg (1995). We need assumptions (A1) and (A2) because a *necessary* condition for the consistency of the IRSGLS estimators of the coefficients of (12) is that the model conditional expectations of  $\zeta_{0it}^{ca}$  and  $\zeta_{1it}^{ca}$ , given  $\hat{Y}_{it}^A = y_{it}^A$  and the coefficient drivers are zero.

## 5.3 Bias- and Error-Corrected Version of the CPS Estimator

Since  $\alpha_{0it}^{CPS} = \varepsilon_{it}^{CPS}$  represents the sum of sampling and non-sampling errors in  $\hat{Y}_{it}^{CPS} = Y_{it} + \varepsilon_{it}^{CPS}$ , satisfying equation (10), it follows that the formula

$$\hat{Y}_{it}^{BECCPS} = \hat{Y}_{it}^{CPS} - \hat{\alpha}_{0it}^{CPS} = \hat{Y}_{it}^{CPS} - \hat{\pi}_{00}^{ca} - \hat{\pi}_{01}^{ca} BP_{it} - \hat{\pi}_{02}^{ca} HP_{it} - \hat{\zeta}_{0it}^{ca} \quad (13)$$

gives “bias- and error-corrected CPS (BECCPS) estimator” of  $Y_{it}$ . A value of this estimator is denoted by  $y_{it}^{BECCPS}$ .

The standard error (SE) of  $\hat{Y}_{it}^{BECCPS}$  is the square root of an approximately unbiased estimator of the model MSE,  $E_m(\hat{Y}_{it}^{BECCPS} - Y_{it})^2$ . Its derivation is given in the Appendix. Cochran’s (1977, p. 14) computations show that the effect of the bias in  $\hat{Y}_{it}^{CPS}$  due to  $u_{it}^{CPS}$  on the probability of the error,  $\hat{Y}_{it}^{CPS} - Y_{it}$ , of more than (or less than) 1.96 (or -1.96) times the standard deviation of  $\hat{Y}_{it}^{CPS}$  is appreciable if the absolute value of the bias is greater than one tenth of the standard deviation of  $\hat{Y}_{it}^{CPS}$ . In these cases, the bias correction in (13) is desirable if the absolute value of the bias remaining in  $\hat{Y}_{it}^{BECCPS}$  after the correction is less than one tenth of the standard deviation of  $\hat{Y}_{it}^{BECCPS}$ . Such desirable bias corrections may frequently occur with  $\hat{Y}_{it}^{BECCPS}$ .

**Applications.** Use the successive difference replication method described in U.S. Department of Labor (2006, Chapter 14) to obtain a new estimate of the within-PSU variance contribution to the total design variance,  $V_p(\hat{Y}_{it}^{CPS} | Y_{it})$ , of  $\hat{Y}_{it}^{CPS}$  in (2). Apply an IRSGLS method to equations (5) and (6) after inserting this new estimate and  $y_{it}^{BECCPS}$  every place an estimate of the within-PSU variance and a CPS estimate of  $Y_{it}$  are used in these equations, respectively. This application can give improved estimates of design effect and  $b_{it}$  used in equation (4).

## 5.4 Improved Additional Estimator

Using  $\hat{\alpha}_{it}^A$  and  $y_{it}^A$  in place of  $\alpha_{it}^A$  and  $\hat{Y}_{it}^A$  used in  $Y_{it} = \alpha_{it}^A \hat{Y}_{it}^A$ , respectively, gives

$$\hat{Y}_{it}^{IA} = \hat{\alpha}_{it}^A y_{it}^A = (\hat{\pi}_{10}^{ca} + \hat{\pi}_{11}^{ca} \frac{BP_{it}}{TP_{it}} + \hat{\pi}_{12}^{ca} \frac{HP_{it}}{TP_{it}} + \hat{\zeta}_{1it}^{ca}) y_{it}^A \quad (14)$$

where  $\hat{Y}_{it}^{IA}$  denotes an improved additional (IA) estimator of  $Y_{it}$ .

The SE of  $\hat{Y}_{it}^{IA}$  is the square root of an approximately unbiased estimator of the model MSE,  $E_m(\hat{Y}_{it}^{IA} - Y_{it})^2$ . Its derivation is given in the Appendix.

### 5.5 Comparison of BECCPS and IA Estimators

A proof of the result,  $\hat{Y}_{it}^{BECCPS} = \hat{Y}_{it}^{IA}$  with probability (w.p.) 1, is given in the Appendix. The following prior information helps us assess the relative accuracies of  $\hat{Y}_{it}^{CPS}$ ,  $\hat{Y}_{it}^A$ ,  $\hat{Y}_{it}^{BECCPS}$ , and  $\hat{Y}_{it}^{IA}$ .

**Prior information.** (i) The overall CPS sample size is sufficient to produce national-level monthly employment or unemployment estimators that satisfy prespecified precision requirements. (ii) However, relatively small CPS state sample sizes do not permit the production of reliable monthly employment and unemployment estimates for the states. (iii) The CPS and CES survey being the household and establishment surveys provide information about “place-of-residence” and “place-of-work” employment, respectively (see U.S. Department of Labor, 1997, p. 45). Consequently,  $\hat{Y}_{it}^A$  may contain a larger magnitude of non-sampling error than  $\hat{Y}_{it}^{CPS}$  when it is viewed as an estimator of “place-of-residence” employment, particularly for metropolitan areas like Washington, DC where a good many suburbanites work (see footnote 12). (iv) Extrapolated QCEW data generated from a time series model may contain large errors because the best nonlinear models that accurately represent asymmetric cycles of state employment are unknown. (v) Both at the national and state levels U.I. Claims are very inaccurate estimates of unemployment.

This prior information implies that at the national level, CPS estimates are more accurate than the corresponding additional estimates; that may or may not be the case for state level data. For this reason,  $\hat{Y}_{it}^{CPS}$  being a state level estimator may not be more precise than  $\hat{Y}_{it}^{BECCPS}$  or  $\hat{Y}_{it}^{IA}$  and  $\hat{Y}_{it}^A$  may not be more precise than  $\hat{Y}_{it}^{CPS}$ . Let  $y_{it}^*$  be an estimate of  $Y_{it}$  based on either  $\hat{Y}_{it}^{BECCPS}$  or  $\hat{Y}_{it}^{IA}$ . Then

**Restriction I.**  $y_{it}^*$  shall lie within an estimated confidence interval of  $y_{it}^{CPS}$ .

Let  $m$  be the number of states or areas which geographically exhaust the entire U.S. Then

**Restriction II.** the sum  $\sum_{i=1}^m y_{it}^*$  shall equal the CPS estimate of the national employment or unemployment,  $\sum_{i=1}^m Y_{it}$ .

The idea of Restriction I is to limit the deviation of  $y_{it}^*$  from  $y_{it}^{CPS}$ . This limiting reduces the loss in efficiency due to the misspecifications in model (12) for some states without losing the benefits of correctly specified model (12) for other states (see Jiang and Lahiri, 2006, p. 36). Prior information (i) stated above justifies Restriction II. The  $y_{it}^*$ 's do not satisfy Restrictions I and II if the coefficient drivers in (10) and (11) are inappropriate. Swamy, Mehta and Chang (2006) found that when  $\hat{Y}_{it}^A$  based on the American Community Survey was used in (9), the coefficient drivers in (10) and (11) were adequate in the sense that they produced the values of  $\hat{Y}_{it}^{BECCPS}$  that satisfied Restriction I for all  $i = 1, 2, \dots, m$ . It is a good practice to experiment with all possible combinations of coefficient drivers on which data are available. After everything that can be done in this way has been done, if some of the resulting  $y_{it}^*$ 's do not satisfy Restriction I, then this restriction may be imposed externally. That is, use  $y_{it}^*$  if  $y_{it}^*$  lies within an estimated confidence interval of  $y_{it}^{CPS}$  and use an estimated lower or upper confidence limit of  $y_{it}^{CPS}$  whichever is closer to  $y_{it}^*$  otherwise (see Rao, 2003, p. 118). Hopefully, major changes in these restricted  $y_{it}^*$ 's are not needed to satisfy Restriction II.

## 6. Time Series of Cross-Sectional Estimates

Estimating model (12) separately for each month denoted by  $t = 1, \dots, \tau$  ( $= \min(\tau_1, \dots, \tau_m)$ ) gives  $(y_{it}^{BECCPS}, a_{0it}^{CPS})'$ ,  $i = 1, \dots, m, t = 1, \dots, \tau$ , where  $y_{it}^{BECCPS}$  and  $a_{0it}^{CPS}$  are the values of  $\hat{Y}_{it}^{BECCPS}$  and  $\hat{\alpha}_{0it}^{CPS}$ , respectively, obtained from cross-sectional data in (12) for each  $t = 1, \dots, \tau$ . To produce seasonally adjusted estimates of the  $Y_{it}$ , the estimates,  $y_{it}^{BECCPS}$ ,  $t = 1, \dots, \tau$ , will be adjusted externally by X-12 ARIMA for each  $i = 1, \dots, m$ . Let  $\hat{\sigma}_{e_{CPS},it}^2$  be an estimate of  $\sigma_{e_{CPS},it}^2$ . Such estimates are computed by the Bureau of Labor Statistics (BLS). An AR (15) model may be fitted to  $a_{0it}^{CPS} / \hat{\sigma}_{e_{CPS},it}$ ,  $t = 1, \dots, \tau$ , for each  $i = 1, \dots, m$ , to estimate the variances and autocovariances of  $e_{it}^{CPS}$  implied by the CPS sampling design.

## 7. Example

In this section,  $Y_{it}$  denotes total employment,  $\hat{Y}_{it}^A = \hat{Y}_{it}^{CES}$ , the CES survey estimator of  $Y_{it}$ ,  $t$  denotes January 2006,  $i$  indexes the 50 United States and the District of Columbia, and  $m = 51$ . Following the IRSGLS method and using the data on the variables in (12) with these values of  $i, t$ , and  $m$ , we obtain the estimates of the coefficients of (12) given below:

$$y_{it}^{CPS} = (0.00000043158 + 0.021745 BP_{it} - 0.008542 HP_{it} + \hat{\zeta}_{0it}^{ca}) \\ + (1.0403 + 0.016003 \frac{BP_{it}}{TP_{it}} - 0.005132 \frac{HP_{it}}{TP_{it}} + \hat{\zeta}_{1it}^{ca}) y_{it}^{CES} \quad (15)$$

(16.99) (0.42495) (-0.18712) (71.028) (0.53134) (-0.18362)

where  $y_{it}^{CES}$  is the CES survey estimate of  $Y_{it}$ , all the other variables are as defined in (12), and the figures in parentheses below the coefficient estimates are the t-ratios.

Equation (15) shows that the estimate, 0.00000043158, of the common national intercept,  $\pi_{00}^{ca}$ , is very close to zero and the estimate, 1.0403, of the common national slope,  $\pi_{10}^{ca}$ , differs slightly from  $(\sum_{i=1}^{51} y_{it}^{CPS} / \sum_{i=1}^{51} y_{it}^{CES}) = 1.0664$ . Both these estimates are significant. The 51 estimates of the errors of (10) and (11) and the coefficients of (9) lie within the ranges:  $-417140 \leq \hat{\zeta}_{0it}^{ca} \leq 207120$ ,  $-0.019683 \leq \hat{\zeta}_{1it}^{ca} \leq 0.0097733$ ,  $-412470 \leq \hat{\alpha}_{0it}^{CPS} \leq 207870$ , and  $1.03 \leq \hat{\alpha}_{1it}^{CES} \leq 1.1238$ . The estimates of the variances of  $\zeta_{0it}^{ca}$  and  $\zeta_{1it}^{ca}$  are:  $9.185E+08$  and  $0.00002045$ , respectively. These estimates show that the coefficient drivers,  $BP_{it}$  and  $HP_{it}$ , included in (10) explain only a very small proportion of the variation in  $\alpha_{0it}^{CPS}$ ,  $i = 1, \dots, 51$ . We need to include additional coefficient drivers in (10) to reduce the variance of  $\zeta_{0it}^{ca}$  to a small number. The capital equipment and the ratio of capital to labor for each area in period  $t$  can serve as additional coefficient drivers. Factors that adjust the “place-of-work” nonfarm employment estimates,  $y_{it}^{CES}$ , to a place-of-residence basis, as in the CPS, can also serve as additional coefficient drivers. Unfortunately, data on these variables are not available.

The CPS estimate,  $y_{it}^{CPS}$ , for each of the 51 areas is given in the column, labeled “CPS”, of Table 1. The values of the BECCPS estimator in (13) for the 51 areas implied by the estimates in (15) are given in the column, labeled “BECCPS”, of Table 1. They are the estimates of the true value component,  $Y_{it}$ , of  $\hat{Y}_{it}^{CPS}$  in (1). The estimate of the error component,  $e_{it}^{CPS} + u_{it}^{CPS}$ , of  $\hat{Y}_{it}^{CPS}$  in (1) for each of the 51 areas is given in the column, labeled “S&NSE”, of Table 1. Note that the estimates in the columns, labeled “BECCPS” and “S&NSE”, are the results of simultaneously estimating the components of  $\hat{Y}_{it}^{CPS}$  using model (12). The coefficient drivers included in equations (10) and (11) are adequate for 36 areas in the sense that the unrestricted BECCPS estimates for these areas satisfy Restriction I. We had to impose Restriction I externally on the BECCPS estimates for the remaining 15 areas.

The two-step estimates of the components,  $Y_{it}$  and  $e_{it}^{CPS}$ , of  $\hat{Y}_{it}^{CPS}$  for each state when  $u_{it}^{CPS}$  is ignored, are given in the columns, labeled, “Signal” and “SurE”, of Table 1, respectively. We have now three different estimates of  $Y_{it}$ . They are denoted by “CPS”, “BECCPS”, and “Signal”, respectively. The CPS estimates are not satisfactory because their accuracy is inadequate, as we have already pointed out in Section 2.1. The Signal estimates are also not satisfactory because they are implied by the incorrect estimates of the parameters of models (2) and (3), as we have already shown in Section 3. Sufficient conditions for the MSE of the BECCPS estimator in (13) to be smaller than the design variance (4) of the CPS estimator,

$\hat{Y}_{it}^{CPS}$ , in (1) are given in the Appendix. These conditions can be satisfied in large samples even when an approximately unbiased estimate of the MSE of (13) is larger than a sample estimate of the design variance in (4), as we have explained in Section 4. If these conditions are not satisfied by model (12) in small samples, then they can be satisfied when we expand the set of coefficient drivers in (10).

The design variance of  $\hat{Y}_{it}^{CPS}$  given in (4) involves unknown quantities and hence is unknown. Its estimate was obtained using  $y_{it}^{CPS}$  and the BLS estimates of  $k_{it}$  and  $D_{e^{CPS},it}$  in place of  $Y_{it}$ ,  $k_{it}$ , and  $D_{e^{CPS},it}$  used in (4), respectively. It is given for each of the 51 areas in the column, labeled “varcps” of Table 1. An alternative estimate of (4) was obtained using the estimate,  $y_{it}^{BECCPS}$ , of  $Y_{it}$  and the IRSGLS estimates of  $\delta_{01}$  and  $\xi_{0itg}$  in place of  $\hat{Y}_{it}^{CPS}(g)$ ,  $\delta_{01}$ , and  $\xi_{0itg}$  used in (5), respectively. It is given for each of the 51 areas in the column, labeled “varbeccps” of Table 1. The “varbeccps” estimates can be more accurate than the “varcps” estimates because the BECCPS estimator in (13) corrects for bias and error in  $\hat{Y}_{it}^{CPS}$ .

Each value given in the column, labeled “VarSignal”, of Table 1 is an estimate of the model variance of the two-step estimator of  $Y_{it}$  for a state. It is an underestimate of the true model variance because the parameters of model (3) are estimated, holding the parameters of model (2) fixed at their incorrectly estimated values.

## 8. Conclusions

Simultaneous estimates of the true value and sampling and non-sampling error components of small area sample estimators presented in this paper are based on weaker assumptions than their two-step estimates. Misspecifications in linking models can result in misestimated components of sample estimators and design variances. Specification errors in the linking models used in simultaneous estimation can be negligible compared to those in the linking models used in two-step estimation.

## Appendix

### A. Derivation of the MSE of the BECCPS Estimator

Let  $x_{it}$  be the 2-vector,  $(1, \hat{Y}_{it}^A)'$ ,  $z_{it}$  be the 5-vector,  $(1, BP_{it}, HP_{it}, \frac{BP_{it}}{TP_{it}}, \frac{HP_{it}}{TP_{it}})'$ ,  $\Pi^{ca}$  be the  $(2 \times 5)$  matrix having  $(\pi_{00}^{ca}, \pi_{01}^{ca}, \pi_{02}^{ca}, 0, 0)$  and  $(\pi_{10}^{ca}, 0, 0, \pi_{11}^{ca}, \pi_{12}^{ca})$  as its first and second rows, respectively, and  $\zeta_{it}^{ca}$  be the 2-vector,  $(\zeta_{0it}^{ca}, \zeta_{1it}^{ca})'$ , where the transpose of a matrix is denoted by a prime. Using these definitions the  $m$  equations in (12) are written

$$\hat{Y}_{it}^{CPS} = X_{zt} \pi^{caLong} + D_{xt} \zeta_{it}^{ca} \quad (A1)$$

where  $\hat{Y}_{it}^{CPS}$  is the  $m$ -vector,  $(\hat{Y}_{1t}^{CPS}, \dots, \hat{Y}_{mt}^{CPS})'$ ,  $X_{zt}$  is the  $(m \times 10)$  matrix having a Kronecker product between  $z_{it}'$  and  $x_{it}'$ , denoted by  $(z_{it}' \otimes x_{it}')$ , as its  $i$ th row,<sup>18</sup>  $\pi^{caLong}$  is the 10-vector given by a column stack of  $\Pi^{ca}$ ,  $D_{xt}$  is the  $(m \times 2m)$  matrix,  $\text{diag}[x_{1t}', \dots, x_{mt}']$ ,  $\zeta_{it}^{ca}$  is the  $2m$ -vector,  $(\zeta_{1t}^{ca}, \dots, \zeta_{mt}^{ca})'$ , and there are zero restrictions on the elements of  $\pi^{caLong}$ . These restrictions can be stated as  $R \pi^{caLong} = 0$ , where  $R$  is the  $4 \times 10$  matrix of full row rank having ones as its (1, 4)-th, (2, 6)-th, (3, 7)-th, and (4, 9)-th elements and zeros elsewhere, and  $0$  is the 4-vector of zeros. Now a  $(6 \times 10)$  matrix  $C$  of full row rank can be found such that  $RC' = 0$ . Under assumptions (10) and (11),  $E_m(D_{xt} \zeta_{it}^{ca} | X_{zt}) = 0$  and  $V_m(D_{xt} \zeta_{it}^{ca} | X_{zt}) = D_{xt}(I_m \otimes \sigma_{\zeta}^2 \Delta_{\zeta})D_{xt}' = \Sigma_{\omega}$ , where  $I$  denotes an identity matrix and a subscript is included to indicate its order, and  $\omega$  is the 4-vector having  $\sigma_{\zeta}^2$  and the distinct elements of  $\Delta_{\zeta}$  as its elements. Swamy and Tinsley (1980) explain how we go from equation (12) to equation (A1).

**Identification in the sense of Lehmann and Casella (1998, p. 24).** The coefficient vector,  $\pi^{caLong}$ , is identifiable if  $X_{zt}$  has full column rank. The error vector,  $\zeta_{it}^{ca}$ , is unidentifiable because  $D_{xt}$  does not have full column rank. This result implies that  $\zeta_{it}^{ca}$  is not consistently estimable (see Lehmann and Casella, 1998, p. 57). The coefficient drivers are used in (10) and (11) to

<sup>18</sup> The definition of a Kronecker product we use is given by Greene (2003, p. 824), among others.

reduce the unidentifiable portions of the coefficients of (9). However,  $D_{xt}\zeta_t^{ca}$  is identifiable and its predictor can be used to obtain a consistent estimator of  $\omega$ .

For known  $\omega$ , applying the derivation in Greene (2003, p. 100) with appropriate modifications to model (A1) gives the MVLUE of  $\pi^{caLong}$  that satisfies the restriction  $R\pi^{caLong} = 0$ . This estimator is

$$\hat{\pi}_R^{caLong}(\omega) = C'(C\Psi_\omega^{-1}C')^{-1}CX'_{zt}\Sigma_\omega^{-1}\hat{Y}_t^{CPS} \quad (A2)$$

where the subscript of  $\hat{\pi}_R^{caLong}$  is shorthand for “restricted” and  $\Psi_\omega = (X'_{zt}\Sigma_\omega^{-1}X_{zt})^{-1}$ . It follows from C.R. Rao (1973, p.77, Problem 33) that  $C'(C\Psi_\omega^{-1}C')^{-1}C = \Psi_\omega - \Psi_\omega R'(R\Psi_\omega R')^{-1}R\Psi_\omega$ . Post-multiplying both sides of this equation by  $\Psi_\omega^{-1}\pi^{caLong}$  gives  $C'(C\Psi_\omega^{-1}C')^{-1}C\Psi_\omega^{-1}\pi^{caLong} = \pi^{caLong}$ , since  $R\pi^{caLong} = 0$ . This result is needed to prove that estimator (A2) is unbiased with the model covariance matrix

$$V_m(\hat{\pi}_R^{caLong}(\omega) | X_{zt}) = C'(C\Psi_\omega^{-1}C')^{-1}C. \quad (A3)$$

For known  $\omega$ , the BLUP of  $\zeta_t^{ca}$  is

$$\hat{\zeta}_{Rt}^{ca}(\omega) = (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D'_{xt} \Sigma_\omega^{-1} M_\omega \hat{Y}_t^{CPS} \quad (A4)$$

where  $M_\omega = I_m - X_{zt} C' (C \Psi_\omega^{-1} C')^{-1} C X'_{zt} \Sigma_\omega^{-1}$ . The matrix,  $M_\omega$ , is idempotent (though not symmetric) with the property that  $M_\omega X_{zt} C' = 0$ . It can be shown that  $E_m(\hat{\zeta}_{Rt}^{ca}(\omega) | X_{zt}) = 0$ ,  $Cov_m[(\hat{\pi}_R^{caLong}(\omega), \hat{\zeta}_{Rt}^{ca}(\omega)) | X_{zt}] = 0$ , and  $V_m(\hat{\zeta}_{Rt}^{ca}(\omega) | X_{zt}) = (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D'_{xt} \Sigma_\omega^{-1} M_\omega D_{xt} (I_m \otimes \sigma_\zeta^2 \Delta_\zeta)$  because  $M_\omega \Sigma_\omega M'_\omega = M_\omega \Sigma_\omega$ .

Pre-multiplying both sides of equation (A4) by  $D_{xt}$  gives  $D_{xt} \hat{\zeta}_{Rt}^{ca}(\omega) = \hat{Y}_t^{CPS} - X_{zt} \hat{\pi}_R^{caLong}(\omega)$  which proves that  $\hat{Y}_{it}^{BECCPS}$  in (13) is equal to  $\hat{Y}_{it}^{IA}$  in (14) with probability 1 for all  $i = 1, \dots, m$  and fixed  $t$ . For known  $\omega$  case, the BECCPS estimator of  $Y_{it}$  in (13) becomes  $\hat{Y}_{it}^{BECCPS}(\omega) = \hat{Y}_{it}^{CPS} - \hat{\alpha}_{0it}^{CPS}(\omega) = \hat{Y}_{it}^{CPS} - (z'_{it} \otimes l'_1) \hat{\pi}_R^{caLong}(\omega) - j'_i D_{l_1} \hat{\zeta}_{Rt}^{ca}(\omega)$ , where  $j_i$  is the  $m$ -vector having 1 as its  $i$ th element and zeros elsewhere,  $l_1$  is the 2-vector,  $(1, 0)'$ , and  $D_{l_1}$  is the  $m \times 2m$  matrix,  $(I_m \otimes l'_1)$ .

The MSE of  $\hat{Y}_{it}^{BECCPS}(\omega)$  is

$$\begin{aligned} E_m[\{\hat{Y}_{it}^{BECCPS}(\omega) - Y_{it}\}^2 | X_{zt}] &= E_m[\{\hat{Y}_{it}^{CPS} - \hat{\alpha}_{0it}^{CPS}(\omega) - \hat{Y}_{it}^{CPS} + \alpha_{0it}^{CPS}\}^2 | X_{zt}] \\ &= g_1(\omega) + g_2(\omega) \end{aligned} \quad (A5)$$

where

$$\begin{aligned} g_1(\omega) &= j'_i D_{l_1} [(I_m \otimes \sigma_\zeta^2 \Delta_\zeta) - (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D'_{xt} \Sigma_\omega^{-1} D_{xt} (I_m \otimes \sigma_\zeta^2 \Delta_\zeta)] D_{l_1} j_i + \\ &[j'_i D_{l_1} (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D'_{xt} \Sigma_\omega^{-1} X_{zt} - 2(z'_{it} \otimes l'_1) C'] C' (C \Psi_\omega^{-1} C')^{-1} C X'_{zt} \Sigma_\omega^{-1} D_{xt} (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D_{l_1} j_i \end{aligned} \quad (A6)$$

and

$$g_2(\omega) = (z'_{it} \otimes l'_1) C' (C \Psi_\omega^{-1} C')^{-1} C (z_{it} \otimes l_1). \quad (A7)$$

As in Rao (2003, p. 99, (6.2.11)), the second term in (A5) arises as a direct consequence of using  $\hat{\pi}_R^{caLong}(\omega)$  instead of  $\pi^{caLong}$  in (A4). In the cases where the coefficient drivers in (10) and (11) reduce the magnitudes of the elements of  $\sigma_\zeta^2 \Delta_\zeta$  to small values, the first term of  $g_1(\omega)$  can be much smaller than Rao’s (2003, pp. 99 and 117)  $g_1(\delta)$  which, in turn, is smaller than the design variance of  $e_{it}^{CPS}$  when  $u_{it}^{CPS}$  is absent. Also,  $g_1(\omega)$  is smaller than its first term if its second term is negative.

Now we can elaborate on our discussion in Section 4. Rao (2003, p. 117) proved that the first term in (A5) is smaller than the design variance of  $e_{it}^{CPS}$  if (i) the non-sampling error,  $u_{it}^{CPS}$ , is zero with probability 1 for all  $i$  and  $t$ , (ii) the sampling errors,  $e_{it}^{CPS}$ ,  $i = 1, \dots, m$ , are independently distributed with known design variances, (iii) for  $i = 1, \dots, m$  and fixed  $t$ , the  $Y_{it}$  follow a GLM model of Rao’s (2003, p. 116) type with no constraints on its coefficients, (iv) the errors of the GLM model (or the  $v_i$ ’s in Rao’s notation) are identically and independently distributed with known model variance, (v) the GLM model error is

independent of  $e_{it}^{CPS}$  for all  $i$  and  $t$ , and (vi) the estimator,  $\hat{Y}_{it}^{BECCPS}$ , is replaced by the BLUP of  $Y_{it}$  given by the GLM model. Even when these conditions hold, the sum of the two terms in (A5) may not be smaller than the design variance of  $e_{it}^{CPS}$  unless  $m$  is sufficiently large and the regularity conditions given in Rao (2003, p. 117) are satisfied. In any case, ignoring the non-sampling error,  $u_{it}^{CPS}$ , can result in an inconsistent and inefficient predictor of  $Y_{it}$ .

We now assume that the error term of model (A1) is normally distributed.

$$E_m(\zeta_t^{ca} | M_\omega \hat{Y}_t^{CPS}) = (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D_{xt}' M_\omega' (M_\omega \Sigma_\omega M_\omega')^{-1} M_\omega \hat{Y}_t^{CPS} \quad (A8)$$

where  $(M_\omega \Sigma_\omega M_\omega')^{-1}$  is a generalized inverse of  $M_\omega \Sigma_\omega M_\omega'$  and this generalized inverse is defined as in C.R. Rao (1973, p. 24). Swamy and Mehta (1975, p. 596) proved that the right-hand side of equation (A8) is equal to the BLUP in (A4). Thus, when  $\zeta_t^{ca}$  is normal, its BLUP is equal to its best unbiased predictor (BUP), since  $E_m(\zeta_t^{ca} | M_\omega \hat{Y}_t^{CPS})$  is the BUP of  $\zeta_t^{ca}$ .

Let  $A$  be a  $m \times (m-6)$  matrix of full column rank such that  $A'X_{zt}C' = 0$ . Then

$$E_m(\zeta_t^{ca} | A' \hat{Y}_t^{CPS}) = (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D_{xt}' A (A' \Sigma_\omega A)^{-1} A' \hat{Y}_t^{CPS}. \quad (A9)$$

It follows from C.R. Rao (1973, p. 77, Problem 33) that  $A(A' \Sigma_\omega A)^{-1} A' + \Sigma_\omega^{-1} X_{zt} C' (C X_{zt}' \Sigma_\omega^{-1} X_{zt} C')^{-1} C X_{zt}' \Sigma_\omega^{-1} = \Sigma_\omega^{-1}$ . Inserting this identity into (A9) shows that (A9) is equal to the BLUP of  $\zeta_t^{ca}$ . This derivation extends Jiang's (1997) proof to the cases where the coefficients of model (A1) are subject to linear restrictions.

We now turn to the case where  $\omega$  is unknown. We use  $A' \hat{Y}_t^{CPS}$  to estimate  $\sigma_\zeta^2 \Delta_\zeta$  so that our estimator, denoted by  $\hat{\sigma}_\zeta^2 \hat{\Delta}_\zeta$ , of  $\sigma_\zeta^2 \Delta_\zeta$  is a function of  $A' \hat{Y}_t^{CPS}$ . Let  $\hat{\omega}$  be the 4-vector having  $\hat{\sigma}_\zeta^2$  and the distinct elements of  $\hat{\Delta}_\zeta$  as its elements. The BECCPS estimator,  $\hat{Y}_{it}^{BECCPS}$ , of  $Y_{it}$  in (13) can be written as  $\hat{Y}_{it}^{BECCPS}(\hat{\omega}) = \hat{Y}_{it}^{CPS} - \hat{\alpha}_{0it}^{CPS}(\hat{\omega}) = \hat{Y}_{it}^{CPS} - (z_{it}' \otimes I_1') \hat{\pi}_R^{caLong}(\hat{\omega}) - j_i' D_{li} \hat{\zeta}_{Ri}^{ca}(\hat{\omega})$ . Sufficient conditions for  $E_m[\hat{Y}_{it}^{BECCPS}(\hat{\omega}) - Y_{it}] = 0$  are given in Kariya and Kurata (2004, pp. 42 and 73). The MSE of  $\hat{Y}_{it}^{BECCPS}(\hat{\omega})$  is

$$\begin{aligned} E_m([\hat{Y}_{it}^{BECCPS}(\hat{\omega}) - Y_{it}]^2 | X_{zt}) &= E_m([\hat{Y}_{it}^{BECCPS}(\omega) - Y_{it}]^2 | X_{zt}) \\ &+ E_m([\hat{Y}_{it}^{BECCPS}(\hat{\omega}) - \hat{Y}_{it}^{BECCPS}(\omega)]^2 | X_{zt}) \\ &+ 2 E_m([\hat{Y}_{it}^{BECCPS}(\omega) - Y_{it}][\hat{Y}_{it}^{BECCPS}(\hat{\omega}) - \hat{Y}_{it}^{BECCPS}(\omega)] | X_{zt}). \end{aligned} \quad (A10)$$

In (A5), we have already evaluated the first term on the right-hand side of this equation. To show that the third term on the right-hand side of this equation vanishes, we first note that  $\hat{Y}_{it}^{BECCPS}(\hat{\omega}) - \hat{Y}_{it}^{BECCPS}(\omega) = -\hat{\alpha}_{0it}^{CPS}(\hat{\omega}) + \hat{\alpha}_{0it}^{CPS}(\omega) = -(z_{it}' \otimes I_1') \hat{\pi}_R^{caLong}(\hat{\omega}) - j_i' D_{li} \hat{\zeta}_{Ri}^{ca}(\hat{\omega}) + (z_{it}' \otimes I_1') \hat{\pi}_R^{caLong}(\omega) + j_i' D_{li} \hat{\zeta}_{Ri}^{ca}(\omega) = (z_{it}' \otimes I_1') \{C'(C\Psi_\omega^{-1}C')^{-1} C X_{zt}' \Sigma_\omega^{-1} - C'(C\Psi_{\hat{\omega}}^{-1}C')^{-1} C X_{zt}' \Sigma_{\hat{\omega}}^{-1}\} [\hat{Y}_t^{CPS} - X_{zt} C'(C\Psi_\omega^{-1}C')^{-1} C \Psi_\omega^{-1} \hat{\pi}_R^{caLong}(\omega)] + j_i' D_{li} \{(I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D_{xt}' \Sigma_\omega^{-1} [\hat{Y}_t^{CPS} - X_{zt} \hat{\pi}_R^{caLong}(\omega)] - (I_m \otimes \hat{\sigma}_\zeta^2 \hat{\Delta}_\zeta) D_{xt}' \Sigma_{\hat{\omega}}^{-1} [\hat{Y}_t^{CPS} - X_{zt} \hat{\pi}_R^{caLong}(\hat{\omega})]\}$  is a function of  $A' \hat{Y}_t^{CPS}$  because  $\hat{\omega}$ ,  $\hat{Y}_t^{CPS} - X_{zt} C'(C\Psi_\omega^{-1}C')^{-1} C \Psi_\omega^{-1} \hat{\pi}_R^{caLong}(\omega)$ ,  $\hat{Y}_t^{CPS} - X_{zt} \hat{\pi}_R^{caLong}(\omega)$ , and  $\hat{Y}_t^{CPS} - X_{zt} \hat{\pi}_R^{caLong}(\hat{\omega})$  are all functions of  $A' \hat{Y}_t^{CPS}$ . Furthermore, the equation,  $\hat{Y}_{it}^{BECCPS}(\omega) - Y_{it} = -(z_{it}' \otimes I_1') [\hat{\pi}_R^{caLong}(\omega) - \pi^{caLong}] - j_i' D_{li} [\hat{\zeta}_{Ri}^{ca}(\omega) - \zeta_t^{ca}]$ , is such that the first term on its right-hand side is independent of  $A' \hat{Y}_t^{CPS}$  because of the condition that  $A'X_{zt}C' = 0$ , and the last term on its right-hand side can be shown to be equal to  $j_i' D_{li} [E_m(\zeta_t^{ca} | A' \hat{Y}_t^{CPS}) - \zeta_t^{ca}]$  using the result in (A9). Hence, the third term on the right-hand side of equation (A10) is equal to  $2 E_m([\hat{Y}_{it}^{BECCPS}(\hat{\omega}) - \hat{Y}_{it}^{BECCPS}(\omega)] E_m[-(z_{it}' \otimes I_1') \{\hat{\pi}_R^{caLong}(\omega) - \pi^{caLong}\} | A' \hat{Y}_t^{CPS}, X_{zt}]) + 2 E_m([\hat{Y}_{it}^{BECCPS}(\hat{\omega}) - \hat{Y}_{it}^{BECCPS}(\omega)] [-j_i' D_{li} \{E_m(\zeta_t^{ca} | A' \hat{Y}_t^{CPS}) - \zeta_t^{ca}\} | A' \hat{Y}_t^{CPS}, X_{zt}])$  which vanishes. This proof extends the proofs of Swamy and Mehta (1969), Jiang (2001), and Rao (2003, p. 114) to the cases where the coefficients of model (A1) are subject to linear restrictions.

Because of the second term on the right-hand side of equation (A10), the MSE of  $\hat{Y}_{it}^{BECCPS}(\hat{\omega})$  is always larger than that of  $\hat{Y}_{it}^{BECCPS}(\omega)$  in the normal case. The method of approximating the MSE of  $\hat{Y}_{it}^{BECCPS}(\hat{\omega})$  by the MSE of  $\hat{Y}_{it}^{BECCPS}(\omega)$  could, therefore, lead to serious underestimation.

Unfortunately, the exact evaluation of the second term on the right-hand side of equation (A10) is generally not possible except in some special cases, as Rao (2003, p. 103) has pointed out. It is therefore necessary to find an approximation to this term. We begin the derivation of such an approximation with the assumptions (see, e.g., Lehmann and Casella, 1998, p. 430) that permit an expansion of  $\hat{\alpha}_{0it}^{CPS}(\hat{\omega})$  about  $\hat{\alpha}_{0it}^{CPS}(\omega)$  with bounded coefficients. Using a Taylor approximation, we obtain

$$\hat{\alpha}_{0it}^{CPS}(\hat{\omega}) - \hat{\alpha}_{0it}^{CPS}(\omega) \approx d(\omega)'(\hat{\omega} - \omega) \quad (A11)$$

where  $d(\omega) = \partial \hat{\alpha}_{0it}^{CPS}(\omega) / \partial \omega$  and it is assumed that the terms involving higher powers of  $\hat{\omega} - \omega$  are of lower order relative to  $d(\omega)'(\hat{\omega} - \omega)$ . Let  $b_1' = j_i' D_{l_1} (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D_{xt}' \Sigma_\omega^{-1}$ . Then  $\hat{\alpha}_{0it}^{CPS}(\omega) = [(z_{it}' \otimes l_1') - b_1' X_{zt}] [\hat{\pi}_R^{caLong}(\omega) - \pi^{caLong}] + (z_{it}' \otimes l_1') \pi^{caLong} + b_1' (\hat{Y}_t^{CPS} - X_{zt} \pi^{caLong})$ . Under normality,

$$d(\omega) \approx (\partial b_1' / \partial \omega) (\hat{Y}_t^{CPS} - X_{zt} \pi^{caLong}) = d^*(\omega), \quad (A12)$$

since the terms involving the derivatives of  $\hat{\pi}_R^{caLong}(\omega) - \pi^{caLong}$  with respect to  $\omega$  are of lower order. Therefore,

$$\begin{aligned} E_m [d(\omega)'(\hat{\omega} - \omega)]^2 &\approx E_m [d^*(\omega)'(\hat{\omega} - \omega)]^2 \approx \text{tr} [E_m (d^*(\omega) d^*(\omega)') \bar{V}(\hat{\omega})] \\ &= \text{tr} [(\partial b_1' / \partial \omega) \Sigma_\omega (\partial b_1' / \partial \omega)' \bar{V}(\hat{\omega})] = g_3(\omega) \end{aligned} \quad (A13)$$

where  $\bar{V}(\hat{\omega})$  is the asymptotic covariance matrix of  $\hat{\omega}$ , and the neglected terms are of lower order. It now follows from (A11)-(A13) that

$$E_m ([\hat{Y}_{it}^{BECCPS}(\hat{\omega}) - \hat{Y}_{it}^{BECCPS}(\omega)]^2 | X_{zt}) = E_m ([\hat{\alpha}_{0it}^{CPS}(\hat{\omega}) - \hat{\alpha}_{0it}^{CPS}(\omega)]^2 | X_{zt}) \approx g_3(\omega). \quad (A14)$$

Inserting (A5) and (A14) into (A10) gives a second-order approximation to the MSE of  $\hat{Y}_{it}^{BECCPS}(\hat{\omega})$  as

$$E_m ([\hat{Y}_{it}^{BECCPS}(\hat{\omega}) - Y_{it}]^2 | X_{zt}) \approx g_1(\omega) + g_2(\omega) + g_3(\omega) \quad (A15)$$

where the terms,  $g_2(\omega)$  and  $g_3(\omega)$ , arise as a direct consequence of using the estimators rather than the true values of  $\pi^{caLong}$  and  $\omega$ , respectively, in  $\hat{Y}_{it}^{BECCPS}(\hat{\omega})$  and are of lower order than  $g_1(\omega)$ . The estimator  $\hat{Y}_{it}^{BECCPS}(\hat{\omega})$  cannot be rejected in favor of  $\hat{Y}_{it}^{CPS}$  even when the design variance of  $e_{it}^{CPS}$  is smaller than the MSE in A(15). The reason is that this design variance understates the MSE of  $\hat{Y}_{it}^{CPS}$  by ignoring  $u_{it}^{CPS}$ .

## B. Estimation of the MSE of the BECCPS Estimator

It follows from Rao (2003, p. 104) that  $E_m g_2(\hat{\omega}) \approx g_2(\omega)$  and  $E_m g_3(\hat{\omega}) \approx g_3(\omega)$  to the desired order of approximation, but  $g_1(\hat{\omega})$  is usually a biased estimator of  $g_1(\omega)$  with a bias that is generally of the same order as  $g_2(\omega)$  and  $g_3(\omega)$ . To evaluate this bias, we make all the assumptions that permit a Taylor expansion of  $g_1(\hat{\omega})$  about  $g_1(\omega)$  with bounded coefficients (see Lehmann and Casella (1998, p. 430)). Under these assumptions,

$$g_1(\hat{\omega}) = g_1(\omega) + (\hat{\omega} - \omega)' \nabla g_1(\omega) + \frac{1}{2} (\hat{\omega} - \omega)' \nabla^2 g_1(\omega) (\hat{\omega} - \omega) \quad (A16)$$

where  $\nabla g_1(\omega)$  is the vector of first-order derivatives of  $g_1(\omega)$  with respect to  $\omega$  and  $\nabla^2 g_1(\omega)$  is the matrix of second-order derivatives of  $g_1(\omega)$  with respect to  $\omega$ . The estimator  $\hat{\omega}$  is generally a biased estimator of  $\omega$  and hence the model expectation of the second term on the right-hand side of equation (A16) is generally nonzero. Consequently,

$$E_m g_1(\hat{\omega}) \approx g_1(\omega) + E_m (\hat{\omega} - \omega)' \nabla g_1(\omega) + \frac{1}{2} \text{tr} [\nabla^2 g_1(\omega) \bar{V}(\hat{\omega})]. \quad (A17)$$

If  $\Sigma_\omega$  has a linear structure, then (A17) reduces to

$$E_m g_1(\hat{\omega}) \approx g_1(\omega) + E_m (\hat{\omega} - \omega)' \nabla g_1(\omega) - g_3(\omega). \quad (A18)$$

This result shows that an estimator of the MSE of  $\hat{Y}_{it}^{BECCPS}(\hat{\omega})$  to the desired order of approximation is given by

$$g_1(\hat{\omega}) - \text{estimate of } [E_m(\hat{\omega} - \omega)' \nabla g_1(\omega)] + g_2(\hat{\omega}) + 2g_3(\hat{\omega}). \quad (\text{A19})$$

The model expectation of (A19) is approximately equal to the MSE of  $\hat{Y}_{it}^{BECCPS}(\hat{\omega})$ . The second term in (A19) can be ignored if it is of lower order than  $-g_3(\omega)$ .

### C. Derivation of the MSE of the IA Estimator

When  $\omega$  is known, the IA estimator of  $Y_{it}$  in (14) can be written as  $\hat{Y}_{it}^{IA}(\omega) = [(z_{it}' \otimes l_2') \hat{\pi}_R^{caLong}(\omega) + j_i' D_{l_2} \hat{\zeta}_{Rt}^{ca}(\omega)] y_{it}^A$ , where  $l_2$  is the 2-vector,  $(0, 1)'$ ,  $j_i$  is the  $m$ -vector having 1 as its  $i$ th element and zeros elsewhere, and  $D_{l_2}$  is the  $(m \times 2m)$  matrix,  $(I_m \otimes l_2')$ . The MSE of  $\hat{Y}_{it}^{IA}(\omega)$  is

$$\begin{aligned} E_m[\{\hat{Y}_{it}^{IA}(\omega) - Y_{it}\}^2 | X_{zt}] &= E_m[\{\hat{\alpha}_{it}^A(\omega) - \alpha_{it}^A\}^2 (y_{it}^A)^2 | X_{zt}] \\ &= f_1(\omega) + f_2(\omega) \end{aligned} \quad (\text{A20})$$

where

$$\begin{aligned} f_1(\omega) &= j_i' D_{l_2} \{(I_m \otimes \sigma_\zeta^2 \Delta_\zeta) - (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D_{xt}' \Sigma_\omega^{-1} D_{xt} (I_m \otimes \sigma_\zeta^2 \Delta_\zeta)\} D_{l_2}' j_i (y_{it}^A)^2 + \{j_i' D_{l_2} \\ &\times (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D_{xt}' \Sigma_\omega^{-1} X_{zt} - 2(z_{it}' \otimes l_2')\} C'(C\Psi_\omega^{-1} C')^{-1} C X_{zt}' \Sigma_\omega^{-1} D_{xt} (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D_{l_2}' j_i (y_{it}^A)^2 \end{aligned} \quad (\text{A21})$$

and

$$f_2(\omega) = (z_{it}' \otimes l_2') C'(C\Psi_\omega^{-1} C')^{-1} C (z_{it}' \otimes l_2) (y_{it}^A)^2. \quad (\text{A22})$$

A second-order approximation to the MSE of  $\hat{Y}_{it}^{IA}(\hat{\omega})$  is

$$E_m([\hat{Y}_{it}^{IA}(\hat{\omega}) - Y_{it}]^2 | X_{zt}) \approx f_1(\omega) + f_2(\omega) + f_3(\omega) \quad (\text{A23})$$

where  $f_3(\omega) = \text{tr}[(\partial b_2' / \partial \omega) \Sigma_\omega (\partial b_2' / \partial \omega)' \bar{V}(\hat{\omega})] (y_{it}^A)^2$  with  $b_2' = j_i' D_{l_2} (I_m \otimes \sigma_\zeta^2 \Delta_\zeta) D_{xt}' \Sigma_\omega^{-1}$ .

### D. Estimation of the MSE of the IA Estimator

An estimator of the MSE of  $\hat{Y}_{it}^{IA}(\hat{\omega})$  to the desired order of approximation is given by

$$f_1(\hat{\omega}) - \text{estimate of } [E_m(\hat{\omega} - \omega)' \nabla f_1(\omega)] + f_2(\hat{\omega}) + 2f_3(\hat{\omega}). \quad (\text{A24})$$

## References

- Brown, L. D. (1990). An Ancillarity Paradox Which Appears in Multiple Linear Regression (with discussion). *The Annals of Statistics*, 18, 471-538.
- Cavanagh, C. L. and Rothenberg, T.J. (1995). Generalized Least Squares with Nonnormal Errors, *Advances in Econometrics and Quantitative Economics*, G.S. Maddala, P.C.B. Phillips and T.N. Srinivasan (eds). Oxford, UK: Blackwell.
- Chang, I., Swamy, P.A.V.B., Hallahan, C. and Tavlas, G.S. (2000). A Computational Approach to Finding Causal Economic Laws. *Computational Economics*, 16, 105-136.
- Cochran, W.G. (1977). *Sampling Techniques*, 3<sup>rd</sup> edition. New York: John Wiley & Sons.
- Durbin, J. and Koopman, S.J. (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.
- Fay, R. and Train, G. (1995). Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for States and Counties. *Proceedings of the Section on Government Statistics*, American Statistical Association.
- Freedman, D.A. and Navidi, W.C. (1986). Regression Models for Adjusting the 1980 Census (with discussion). *Statistical Science*, 1, 3-39.
- Greene, W.H. (2003). *Econometric Analysis*, 5<sup>th</sup> edition. Upper Saddle River, New Jersey: Prentice Hall.
- Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Havenner, A. and Swamy, P.A.V.B. (1981). A Random Coefficient Approach to Seasonal Adjustment of Economic Time Series. *Journal of Econometrics*, 15, 177-209.
- Hwang, J.S. and Dempster, A.P. (1999). A Stochastic System for Modeling Labor Force Series of Small Areas. *Statistica Sinica*, 9, 297-324.
- Jiang, J. (1997). A Derivation of BLUP – Best Linear Unbiased Predictor. *Statistics & Probability Letters*, 32, 321-324.
- Jiang, J. (2001). Private Communication to J.N.K. Rao.
- Jiang, J. and Lahiri, P. (2006). Mixed Model Prediction and Small Area Estimation (with discussion). *Sociedad de Estadística e Investigación Operativa Test*, 15, 1-96.
- Kariya, T. and Kurata, H. (2004). *Generalized Least Squares*. Hoboken, New Jersey: John Wiley & Sons, Ltd.
- Lehmann, E.L. (1999). *Elements of Large-Sample Theory*, New York: Springer.
- Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation*. 2<sup>nd</sup> edition. New York: Springer.
- Lent, J. (1991). Variance Estimation for Current Population Survey Small Area Labor Force Estimates. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.
- Lent, J. (1994). Memorandum for Michael Welch, Bureau of Labor Statistics, Washington, DC.
- Little, R.J. (2004). To Model or Not To Model? Competing Modes of Inference for Finite Population Sampling. *Journal of the American Statistical Association*, 99, 546-556.
- Maddala, G.S. and Kim, I. (1998). *Unit Roots, Cointegration, and Structural Change*. Cambridge: Cambridge University Press.

- Magnac, T. and Visser, M. (1999). Transition Models with Measurement Errors. *The Review of Economics and Statistics*, 81, 466-474.
- Mathiowetz, N.A. and Duncan, G.J. (1988). Out of Work, Out of Mind: Response Errors in Retrospective Reports of Unemployment. *Journal of Business & Economic Statistics*, 6, 221-229.
- Memorandum from Khandaker Mansur. "Documentation of the CPS History of Design Effects, Sampling Intervals, and Type A Rates. Procedures and Sources Used to Obtain These Estimates (Var2000-3)," dated February 22, 2007.
- Montgomery, A.L., Zarnowitz, V., Tsay, R.S. and Tiao, G.C. (1998). Forecasting the U.S. Unemployment Rate. *Journal of the American Statistical Association*, 93, 478-493.
- Pfeffermann, P. and Tiller, R. (2006). Small-Area Estimation with State-Space Models Subject to Benchmark Constraints. *Journal of the American Statistical Association*, 101, 1387-1397.
- Rao, C.R. (1973). *Linear Statistical Inferences and Its Applications*, 2<sup>nd</sup> edition. New York: John Wiley & Sons.
- Rao, J.N.K. (2003). *Small Area Models*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Rowthorn, R. and Glyn, A.J. (2006). Convergence and Stability in U.S. Employment Rates. *Contributions to Macroeconomics*, 6, 1-42.
- Swamy, P.A.V.B. and Mehta, J.S. (1969). On Theil's Mixed Regression Estimator. *Journal of the American Statistical Association*, 64, 273-276.
- Swamy, P.A.V.B. and Mehta, J.S. (1975). Bayesian and Non-Bayesian Analysis of Switching Regressions and of Random Coefficient Regression Models. *Journal of the American Statistical Association*, 70, 593-602.
- Swamy, P.A.V.B. and Tinsley, P.A. (1980). Linear Prediction and Estimation Methods for Regression Models with Stationary Stochastic Coefficients. *Journal of Econometrics*, 12, 103-142.
- Swamy, P.A.V.B., Zimmerman, T.S., Mehta, J.S. and Robison, E.L. (2005). New Improved Small Area Models. Federal Committee on Statistical Methodology Research Conference (Conference Papers).
- Swamy, P.A.V.B., Yaghi, W., Mehta, J.S. and Chang, I. (2007). Empirical Best Linear Unbiased Prediction in Misspecified and Improved Panel Data Models with an Application to Gasoline Demand. *Computational Statistics & Data Analysis*, 51, 3381-3392.
- Swamy, P.A.V.B., Mehta, J.S. and Chang, I. (2006). An Efficient Method of Estimating the True Value of a Population Characteristic from its Discrepant Estimates. Unpublished manuscript.
- Tiller, R.B. (1992). Time Series Modeling of Sample Survey Data from the U.S. Current Population Survey. *Journal of Official Statistics*, 8, 149-166.
- Tiller, R.B. (2005). State Model-Based Labor Force Estimation. Unpublished manuscript.
- Wolter, K.M. (1985). *Introduction to Variance Estimation*. New York: Springer.
- Zimmerman, T.S. and Robison, E.L. (1995). Estimation of Autocorrelations for Current Population Survey Labor Force Characteristics. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 414-419.
- Zimmerman, T.S. (2005). Reliability of CPS Estimators Under Current (2000) Design. Memorandum.
- Zimmerman, T.S. (2007). Sampling Error Correlation Status Report. Memorandum.

U.S. Department of Labor, Bureau of Labor Statistics (1997). BLS Handbook of Methods, Bulletin 2490. Washington, DC: Government Printing Office.

U.S. Department of Labor, Bureau of Labor Statistics (2006). Design and Methodology: Current Population Survey, Technical Paper 66. Washington, DC: Government Printing Office.

Table 1. Two-Step versus Simultaneous Estimates of True Value and Error Components of Sample Estimates of State Employment for January 2006

St	CPS Sample Estimates		Simultaneous Estimates			Two-Step Estimates		
	CPS	varcps	BECCPS	S&NSE	varbeccps	Signal	VarSignal	SurE
AL	2038673	3679861578	2022994	15679	3679900196	2064406	1677393936	-25733
AK	308055	95009042	304142	3913	95009433	312698	34951744	-4643
AZ	2719607	4940016080	2710402	9206	4940037094	2742883	1872639076	-23276
AR	1335773	931402110	1275956	59817	931377568	1279302	386712225	56471
CA	16568973	19713973309	16565410	3563	19713785649	16815860	19757394721	-246887
CO	2417786	2594841766	2339775	78011	2594846962	2428008	1521468036	-10222
CT	1728484	1184483549	1722675	5809	1184456506	1709671	577200625	18813
DE	430941	71573417	441423	-10482	71573611	416387	27405225	14554
DC	274617	37453913	286613	-11996	37453577	292316	25030009	-17699
FL	8524044	15431076249	8506356	17688	15431008706	8415622	7395140025	108422
GA	4418295	5567633630	4272046	146249	5567646053	4399394	2121523600	18901
HI	637847	136983872	627152	10695	136983835	626053	78021889	11794
ID	723132	314448720	688376	34756	314459782	701278	118766404	21854
IL	6102978	8049548107	6106455	-3477	8049551901	6074746	4563948249	28232
IN	3116148	4327930599	3050431	65717	4327891224	3051273	2294601604	64875
IA	1588221	811523654	1532386	55835	811511184	1553633	377408329	34588
KS	1400258	822429885	1375147	25111	822416077	1386284	295083684	13974
KY	1811409	2988915564	1881199	-69790	2988889434	1856101	896403600	-44692
LA	1862668	2733058634	1858502	4166	2733026297	1869696	1643572681	-7028
ME	666650	186489439	639884	26766	186490357	660612	88717561	6038
MD	2844012	2088972084	2754430	89582	2088966050	2817955	929579121	26057
MA	3117108	5040669842	3256264	-139156	5040668540	3182095	1567684836	-64987
MI	4664360	7475499516	4501730	162630	7475551167	4663248	3227944225	1112
MN	2714723	2432139574	2810682	-95959	2432132934	2765644	1109889225	-50921
MS	1205433	1491094340	1154127	51306	1491091754	1195684	725709721	9749
MO	2858183	4227831478	2827705	30478	4227860017	2831368	2008832400	26815
MT	470794	169403823	445284	25510	169404383	458981	45131524	11813
NE	912968	256986524	944388	-31420	256987727	925837	120253156	-12869
NV	1172640	694617481	1224297	-51657	694624454	1176211	238702500	-3571
NH	712864	131084819	690423	22441	131084294	703637	54730404	9227
NJ	4186935	4993905767	4199985	-13050	4993927587	4226153	2227462416	-39218
NM	890814	616281651	854395	36419	616295929	877161	110250000	13653
NY	9051645	16731126410	9024711	26934	16731098425	8935803	10327437376	115842
NC	4125019	6143432242	4111840	13179	6143465977	4117926	2625537600	7093
ND	350767	70109708	355857	-5090	70109997	332271	32821441	18496
OH	5464068	10148532036	5561621	-97553	10148624788	5447573	2539353664	16495
OK	1647282	1897191571	1585666	61616	1897214340	1628011	651985156	19271
OR	1747033	1377585050	1732349	14684	1377595302	1737507	462121009	9526
PA	5905700	6871837932	5902957	2743	6871921628	5879965	3680484889	25735
RI	543188	89395889	524657	18531	89395489	540791	54405376	2397
SC	1932052	2196475459	1931836	216	2196491622	1922926	548074921	9126
SD	403813	70320546	400649	3164	70320508	401614	26081449	2199
TN	2799013	4799435890	2840960	-41947	4799471524	2727933	1528731801	71080
TX	10675536	22253332600	10753712	-78176	22253467249	10647979	7967704644	27557
UT	1213992	810295800	1216001	-2010	810306676	1214375	356643225	-383
VT	353711	40295008	341269	12442	40295305	340964	18198756	12747
VA	3737584	5901222652	3816349	-78765	5901175918	3793403	1311453796	-55819
WA	3143887	2853273738	3039191	104696	2853224390	3110833	1535307489	33054
WV	777019	700320633	769822	7197	700338137	747780	268304400	29239
WI	2914222	3758736287	2915293	-1072	3758743232	2845595	1782106225	68627
WY	270567	42070126	273908	-3341	42069787	265398	15856324	5169
Tot	1411481487		140969680	511806		141118844		362643
				-511806				-362643

Note: CPS =  $y_{it}^{CPS}$ , varcps =  $\hat{V}_p(\hat{Y}_{it}^{CPS} | Y_{it})$  in Section 3.1, BECCPS =  $y_{it}^{BECCPS}$  in Section 5.3, S&NSE = Estimate of  $\epsilon_{it}^{CPS}$  in Section 5.2, varbeccps =  $\hat{\sigma}_{\epsilon_{it}^{CPS}}^2$  in Section 7, Signal = Two-step estimate of  $Y_{it}$ , VarSignal = estimated model variance of Signal, SurE = Two-step estimate of  $e_{it}^{CPS}$ .